BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/12, C07K 14/47, C12N 1/21, 5/10, 9/16, A61K 38/17, 39/395, 38/46, 31/70, G01N 33/68, C12Q 1/44, 1/42

A1

(11) International Publication Number: **WO 98/37196**

(43) International Publication Date: 27 August 1998 (27.08.98)

(21) International Application Number: PCT/US98/03323

(22) International Filing Date: 19 February 1998 (19.02.98)

(30) Priority Data:
08/805,583 25 February 1997 (25.02.97) US

(71) Applicant: LUDWIG INSTITUTE FOR CANCER RESEARCH [CH/US]; 1345 Avenue of the Americas, New York, NY 10105 (US).

(72) Inventors: SARAS, Jan; Lindbergsgatan 15B, S–752 40 Uppsala (SE). FRANZEN, Petra; Lindbergsgatan 15B, S–752 40 Uppsala (SE). ASPENSTROM, Pontus; Salagatan 30A, S–753 30 Uppsala (SE). HELLMAN, Ulf; Borjegatan 40, S–752 29 Uppsala (SE). GONEZ, Leonel, Jorge; 4/47–49 Willesden Road, Hughesdale, VIC 3166 (AU). HELDIN, Carl–Henrik; Hesselmans vag 35, S–752 63 Uppsala (SE).

(74) Agent: VAN AMSTERDAM, John, R.; Wolf, Greenfield & Sacks, P.C., 600 Atlantic Avenue, Boston, MA 02210 (US).

(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: PARG, A GTPase ACTIVATING PROTEIN WHICH INTERACTS WITH PTPL1

(57) Abstract

The invention describes nucleic acids encoding the PARG protein, including fragments and biologically functional variants thereof. Also included are polypeptides and fragments thereof encoded by such nucleic acids, and antibodies relating thereto. Methods and products for using such nucleic acids and polypeptides also are provided.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

PARG, A GTPase ACTIVATING PROTEIN WHICH INTERACTS WITH PTPL1

Field of the Invention

This invention relates to nucleic acids and encoded polypeptides which interact with the PTPL1 phosphatase and which are GTPase activating proteins. The invention also relates to agents which bind the nucleic acids or polypeptides. The invention further relates to methods of using such nucleic acids and polypeptides in the treatment and/or diagnosis of disease.

Background of the Invention

The Rho family of Ras-like GTPases, which includes Rho, Rac and Cdc42, control actin-based cytoskeletal rearrangements (reviewed in Hall, *Annu. Rev. Cell Biol.* 10:31-54, 1994; Zigmond, *Curr. Opin. Cell Biol.* 8:66-73, 1996). Rho regulates receptor-mediated assembly of focal adhesions and stress fibers (Ridley and Hall, *Cell* 70:389-399, 1992), while Rac regulates the formation of membrane ruffles (Ridley *et al.*, *Cell* 70:401-410, 1992) and Cdc42 controls the formation of filopodia (Nobes and Hall, *Cell* 81:53-62, 1995). Rho proteins have also been shown to be important in the regulation of cell proliferation (reviewed in Symons, *Trends Biochem. Sci.* 21:178-181, 1996). As members of the Ras superfamily, Rho proteins function as molecular switches, having an active, GTP-bound form, and an inactive, GDP-bound form. The active, GTP-bound form, is negatively regulated by GTPase activating proteins (GAPs) which enhance the intrinsic GTPase activity of Rho proteins. A number of GAPs that are active on proteins of the Rho family have been identified (reviewed in Lamarche and Hall, *TIG* 10:436-440, 1994). These include p50RhoGAP (Lancaster *et al.*, *J. Biol. Chem.* 269:1137-1142, 1994), Myr5 (Reinhard *et al.*, *EMBO J.* 14:697-704, 1995), and p190 (Settleman *et al.*, *Nature* 359:153-154, 1992) which are also active on Rac and Cdc42. Another GAP, p122-RhoGAP (Homma and Emori, *EMBO J.* 14:286-291, 1995) appears to be specific for Rho.

Intracellular protein tyrosine phosphatases (PTPs) are a diverse group of proteins involved in signal transduction (reviewed in Streuli, *Curr. Opin. Cell Biol.* 8:182-188, 1996). They contain a conserved PTP domain which specifically dephosphorylates tyrosine residues and, in addition, domains that regulate their subcellular localization and activity (reviewed in Mauro and Dixon, *Trends Biochem. Sci.* 19:151-155, 1994). For example, the SH2 domains of SHP-1 and SHP-2 enables these PTPs to localize to and interact with activated growth factor receptors (Mauro and Dixon, 1994). Correct localization of PTPs is of importance, since the

PTP domains usually have broad substrate specificity.

PTPL1 (Saras *et al.*, *J. Biol. Chem.* 269:24082-24089, 1994) also called PTP-BAS (Maekawa *et al.*, *FEBS Lett.* 337:200-206, 1994), hPTP1E (Banville *et al.*, *J. Biol. Chem.* 269:22320-22327, 1994) and FAP-1 (Sato *et al.*, *Science* 268:411-415, 1995), is a 250 kDa protein expressed in many tissues and cell lines. PTPL1 is fully described in PCT published application WO95/06735. It contains an N-terminal leucine zipper motif followed by a domain with homology to the Band 4.1 superfamily. Band 4.1-like domains are found in proteins involved in the linkage of actin filaments to the plasma membrane (Arpin *et al.*, *Curr. Opin. Cell Biol.* 6:136-141, 1994). Five PDZ domains [PDZ is derived from PSD-95 (Cho *et al.*, *Neuron* 9:929-942, 1992), Dlg-A (Woods and Bryant, *Cell* 66:451-464, 1991) and ZO-1 (Itoh *et al.*, *J. Cell. Biol.* 121:491-502, 1993), each of which contains three such domains] are present between the Band 4.1-like domain and the C-terminal PTP domain. These domain structures of about 90 amino acid residues have also been called GLGF repeats or DHRs and are identified in a variety of proteins (Ponting and Phillips, *Trends Biochem. Sci.* 20:102-103, 1995). A PDZ domain of PTPL1 has been shown to interact with the C-terminal tail of the membrane receptor Fas (Sato *et al.*, 1995) and PDZ domains of PSD-95 bind to the C-terminals of the NMDA-receptor and Shaker-type K^+ channels (Kim *et al.*, *Nature* 378:85-88, 1995; Kornau *et al.*, *Science* 269:1737-1740, 1995). The crystal structures of two PDZ domains have recently been published (Doyle *et al.*, *Cell* 85:1067-1076, 1996; Morais Cabral *et al.*, *Nature* 382:649-652, 1996).

There exists a need to influence the receptor-mediated intracellular signal transduction pathways to treat disease. There also exists a need to identify the gene(s) responsible for increased or decreased signal transduction and to provide a genetic therapy for treating diseases resulting from aberrant signal transduction.

An object of the invention is to provide compounds that desirably influence the signal transduction by the Rho family of Ras-like GTPases.

Another object of the invention is to provide therapeutics for treating diseases resulting from aberrant signal transduction by the Rho family of Ras-like GTPases.

Still another object of the invention is to provide diagnostics and research tools relating to PARG, PTPL1 and the Rho family of Ras-like GTPases. These and other objects will be described in greater detail below.

Summary of the Invention

The invention provides isolated nucleic acid molecules, unique fragments of those molecules, expression vectors containing the foregoing, and host cells transfected with those molecules. The invention also provides isolated polypeptides and agents which bind such polypeptides, including antibodies. The foregoing can be used in the diagnosis or treatment of conditions characterized by the expression of a PARG nucleic acid or polypeptide. The invention also provides methods for identifying pharmacological agents useful in the diagnosis or treatment of such conditions. Here, we present the cDNA cloning of a PTPL1-associated RhoGAP, PARG, a 150 kDa protein that contains a GAP domain that displays strong activity towards Rho. Furthermore, the C-terminal tail of PARG specifically interacts with the fourth PDZ domain (PDZ4) of PTPL1.

According to one aspect of the invention, an isolated nucleic acid molecule is provided. The molecule hybridizes under stringent conditions to a molecule consisting of the nucleic acid sequence of SEQ ID NO:1. The isolated nucleic acid molecule codes for a GTPase activating polypeptide. The invention further embraces nucleic acid molecules that differ from the foregoing isolated nucleic acid molecules in codon sequence due to the degeneracy of the genetic code. The invention also embraces complements of the foregoing nucleic acids.

In preferred embodiments, the isolated nucleic acid molecule comprises a molecule consisting of the nucleic acid sequence of SEQ ID NO:1. More preferably, the isolated nucleic acid molecule comprises a molecule consisting of nucleotides 184-3966 of SEQ ID NO:1. Preferably the isolated nucleic acid comprises a molecule having a sequence which encodes amino acids 666-853 of SEQ ID NO:2, amino acids 613-652 of SEQ ID NO:2, and/or amino acids 193-509 of SEQ ID NO:2.

According to another aspect of the invention, an isolated nucleic acid molecule is provided. The isolated nucleic acid molecule comprises a molecule consisting of a unique fragment of nucleotides 184-3966 of SEQ ID NO:1 between 12 and 3781 nucleotides in length and complements thereof, provided that the isolated nucleic acid molecule excludes sequences consisting only of SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5 and SEQ ID NO:13. In one embodiment, the isolated nucleic acid molecule consists of between 12 and 32 contiguous nucleotides of SEQ ID NO:1, or complements of such nucleic acid molecules. In preferred embodiments, the unique fragment is at least 14, 15, 16, 17, 18, 20 or 22 contiguous nucleotides of the nucleic acid sequence of SEQ ID NO:1, or complements thereof.

According to another aspect of the invention, an isolated nucleic acid molecule which encodes a PDZ domain binding site is provided, comprising a sequence selected from the group consisting of SEQ ID NO:6, SEQ ID NO:8 and SEQ ID NO:10, or nucleic acid molecules that differ from the nucleic acid molecules of the group consisting of SEQ ID NO:6, SEQ ID NO:8 and SEQ ID NO:10 in codon sequence due to the degeneracy of the genetic code. Preferably the isolated nucleic acid consists of a molecule having a sequence selected from the group consisting of SEQ ID NO:6, SEQ ID NO:8 and SEQ ID NO:10.

According to another aspect of the invention, the invention involves expression vectors, and host cells transformed or transfected with such expression vectors, comprising the nucleic acid molecules described above.

According to another aspect of the invention, an isolated polypeptide is provided. The isolated polypeptide is encoded by the isolated nucleic acid molecule of claim 1, 2 or 14, and the polypeptide has GTPase activating activity. In preferred embodiments, the isolated polypeptide comprises a polypeptide having the sequence of amino acids 658-898 of SEQ ID NO:2.

According to a further aspect of the invention, an isolated polypeptide is provided. The isolated polypeptide comprises a polypeptide encoded by a nucleic acid which hybridizes under stringent conditions to nucleotides 2020-2139 of SEQ ID NO:1. In preferred embodiments, the isolated polypeptide comprises a polypeptide having the sequence of amino acids 613-652 of SEQ ID NO:2 is provided. The isolated polypeptide has a Cys-rich domain.

According to another aspect of the invention, an isolated polypeptide is provided. The isolated polypeptide comprises a polypeptide encoded by a nucleic acid which hybridizes under stringent conditions to nucleotides 760-1710 of SEQ ID NO:1. In preferred embodiments, the isolated polypeptide comprises a polypeptide having the sequence of amino acid 193-509 of SEQ ID NO:2 is provided. The isolated polypeptide is a ZPH domain polypeptide.

In other embodiments, the isolated polypeptide consists of a fragment or variant of the foregoing which retains the activity of the foregoing.

According to still another aspect of the invention, an isolated polypeptide is provided. The isolated polypeptide is encoded by a nucleic acid molecule having a sequence selected from the group consisting of SEQ ID NO:6, SEQ ID NO:8 and SEQ ID NO:10. The isolated polypeptide comprises a polypeptide selected from the group consisting of a polypeptide having the sequence of SEQ ID NO:7, a polypeptide having the sequence of SEQ ID NO:9, and a polypeptide having the sequence of SEQ ID NO:11.

According to another aspect of the invention, there are provided isolated polypeptides which selectively bind a PARG protein or fragment thereof. The isolated polypeptide in certain embodiments binds to a polypeptide comprising the sequence of amino acids 658-898 of SEQ ID NO:2, amino acids 613-652 of SEQ ID NO:2, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11 or amino acids 193-509 of SEQ ID NO:2. The isolated polypeptide preferably binds to a polypeptide consisting essentially of the sequence of amino acids 658-898 of SEQ ID NO:2, amino acids 613-652 of SEQ ID NO:2, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11 or amino acids 193-509 of SEQ ID NO:2. In preferred embodiments, isolated binding polypeptides include antibodies and fragments of antibodies (e.g., Fab, $F(ab)_2$, Fd and antibody fragments which include a CDR3 region which binds selectively to the PARG polypeptides of the invention).

The invention provides in another aspect an isolated complex of polypeptides. The isolated complex includes a PTPL1 polypeptide, such a polypeptide including the amino acid sequence of SEQ ID NO:12 bound to a polypeptide as claimed in claim 1. The isolated complex has both PTPL1 phosphatase activity and PARG GAP activity. Preferably the isolated complex consists essentially of the polypeptide of SEQ ID NO:12 and the polypeptide of SEQ ID NO:2.

According to still another aspect of the invention, methods for reducing Rho family GTPase signal transduction in a mammalian cell are provided. The methods involve administering to a mammalian cell an amount of an inhibitor of Rho family GTPase activity effective to reduce Rho family GTPase signal transduction in the mammalian cell. In certain embodiments, the inhibitor is an isolated PARG polypeptide, having Rho GAP activity, encoded by SEQ ID NO:1. In other embodiments, the inhibitor is an isolated complex of polypeptides comprising a polypeptide comprising the amino acid sequence of SEQ ID NO:12 and a polypeptide comprising the amino acid sequence of SEQ ID NO:2.

According to still another aspect of the invention, methods for reducing proliferation of a cancer cell are provided. The methods involve administering to a cancer cell an amount of a PARG polypeptide, comprising a polypeptide encoded by the nucleic acid of claim 1, effective to reduce proliferation of the cancer cell.

The invention in a further aspect provides methods for increasing Rho family GTPase signal transduction in a mammalian cell. A dominant negative variant of the polypeptide of SEQ ID NO:2 is administered to the mammalian cell in an amount effective to increase Rho family GTPase signal transduction. Preferably the dominant negative polypeptide includes an

inactivated GTPase activating domain which contains a deletion or at least one inactivating point mutation.

According to a further aspect of the invention, methods for reducing binding of a protein which includes a PDZ4 domain to a protein which includes a PDZ4 domain binding site are provided. The methods involve contacting the protein which includes PDZ4 domain with an agent which binds to the PDZ4 domain for a time effective to reduce the binding of the protein which includes PDZ4 domain to the protein which includes PDZ4 domain binding site. In certain embodiments the agent is an isolated peptide and includes at its carboxyl terminus the amino acid sequence of SEQ ID NO:7. The isolated peptide can include conservative substitutions of the amino acid sequence of SEQ ID NO:7, excepting the terminal valine. In preferred embodiments the amino acid sequence of the peptide is selected from the group consisting of SEQ ID NO:7, SEQ ID NO:9 and SEQ ID NO:11. In other embodiments the agent is an antibody which binds to the PDZ4 domain, preferably a monoclonal antibody. In some embodiments, methods provide inhibiting binding of a protein which includes a PDZ4domain and a protein which includes a PDZ4domain binding site in a mammalian cell. Such methods involve contacting the mammalian cell with an agent which binds to the PDZ4 domain for a time effective to reduce the binding of the protein which includes PDZ4 domain to the protein which includes PDZ4 domain binding site.

The invention in another aspect provides methods of modulating mast cell secretion in a subject. The methods include administering to the subject in need of such treatment an amount of a modulator of PARG GTPase activating activity effective to modulate mast cell secretion in the subject.

The invention in still another aspect provides compositions comprising a PARG polypeptide which has GTPase activating activity, a complex of such a PARG polypeptide and PTPL1 phosphatase, or a peptide agent which binds to a PDZ4 domain and which includes the sequence of SEQ ID NO:7, and a pharmaceutically acceptable carrier.

The invention in a further aspect involves a method for decreasing PARG GTPase activating activity in a subject. An agent that selectively binds to an isolated nucleic acid molecule of the invention or an expression product thereof is administered to a subject in need of such treatment, in an amount effective to decrease PARG GTPase activating activity in the subject. Preferred agents are antisense nucleic acids, including modified nucleic acids, and polypeptides.

According to another aspect of the invention, methods are provided for identifying lead compounds for a pharmacological agent useful in the diagnosis or treatment of disease associated with PARG GTPase activating activity or with PARG binding to a protein containing a PDZ4 domain. The methods involve forming a mixture of a PARG polypeptide or fragment thereof containing a GTPase activating domain or a PDZ4 domain binding site, a protein which interacts with the foregoing GTPase activating domain or PDZ4 domain binding site, and a candidate pharmacological agent. The mixture is incubated under conditions which, in the absence of the candidate pharmacological agent, permit a first amount of specific activation of the GTPase by the PARG GTPase activating domain or permit a first amount of selective binding of the protein containing a PDZ4 domain by the PDZ4 domain binding site. A test amount of the specific activation of the GTPase by the PARG GTPase activating domain or the selective binding of the protein containing a PDZ4 domain by the PDZ4 domain binding site then is detected. Detection of an increase in the foregoing activities in the presence of the candidate pharmacological agent indicates that the candidate pharmacological agent is a lead compound for a pharmacological agent which increases specific activation of the GTPase by the PARG GTPase activating domain or selective binding of the protein containing a PDZ4 domain by the PDZ4 domain binding site. Detection of a decrease in the foregoing activities in the presence of the candidate pharmacological agent indicates that the candidate pharmacological agent is a lead compound for a pharmacological agent which decreases specific activation of the GTPase by the PARG GTPase activating domain or selective binding of the protein containing a PDZ4 domain by the PDZ4 domain binding site. Where the activity tested is specific activation of the GTPase, the protein which interacts with the GTPase activating domain preferably is Rho. Where the activity tested is selective binding of a PDZ4 domain, the protein which interacts with the PDZ4 domain binding site preferably is PTPL1.

The invention also contemplates specifically the use of the foregoing compositions in the manufacture of a medicament, particularly medicaments for treating conditions characterized by aberrant Rho family protein signal transduction, cell proliferation and/or mast cell secretion.

These and other objects of the invention will be described in further detail in connection with the detailed description of the invention.

Brief Description of the Drawings

Figure 1 is a representation of the production GST-PDZ fusion proteins. (A) Schematic

illustration of the GST-PDZ fusion proteins showing the domain structure of PTPL1 and the design of PTPL1 -derived GST-PDZ fusion proteins (B) Expression of GST-PDZ fusion proteins.

Figure 2 shows the interaction of GST-PDZ fusion proteins with components in cell lysate.

Figure 3 depicts the structure of PARG protein. (A) Deduced amino acid sequence of PARG. (B) Comparison of amino acid sequences of ZPH regions found in PARG and in the gene product of the *C. elegans* gene ZK669.1a. (C) Schematic diagram illustrating the domain structure of PARG and ZK669.1a.

Figure 4 shows Northern blot analysis of expression of PARG mRNA in different human tissues.

Figure 5 shows an analysis of the GAP activity of PARG. **(A)** Expression of the GAP domain of PARG as a GST fusion protein. Rho **(B)**, Rac **(C)**, and Cdc42 **(D)** loaded with γ-^{32}P-GTP were incubated with 1 nM (open circles), 20 nM (filled circles) of the GAP domain of PARG expressed as a GST fusion protein, or 1 00 nM GST (squares) as a control, for different time periods at 30°C.

Figure 6 shows binding of GST-PDZ fusion proteins to a C-terminal PARG peptide.

Brief Description of the Sequences

SEQ ID NO:1 is the nucleotide sequence of the PARG cDNA.

SEQ ID NO:2 is the amino acid sequence of the translation product of the PARG cDNA, including a RhoGAP domain at amino acids 666-853, a cysteine-rich domain at amino acids 613-652, a ZPH domain at amino acids 193-509 of SEQ ID NO:2, and a carboxyl-terminal PDZ domain binding site.

SEQ ID NO:3 is the nucleotide sequence of the expressed sequence tag identified by GenBank accession number T32345.

SEQ ID NO:4 is the nucleotide sequence of the expressed sequence tag identified by GenBank accession number Z28937.

SEQ ID NO:5 is the nucleotide sequence of the expressed sequence tag identified by GenBank accession number Z28520.

SEQ ID NO:6 is the nucleotide sequence encoding the PARG PDZ domain binding site which consists of 4 amino acids.

SEQ ID NO:7 is the amino acid sequence of the PARG PDZ domain binding site which consists of 4 amino acids.

SEQ ID NO:8 is the nucleotide sequence encoding the PARG PDZ domain binding site which consists of 5 amino acids.

SEQ ID NO:9 is the amino acid sequence of the PARG PDZ domain binding site which consists of 5 amino acids.

SEQ ID NO:10 is the nucleotide sequence encoding the PARG PDZ domain binding site which consists of 6 amino acids.

SEQ ID NO:11 is the amino acid sequence of the PARG PDZ domain binding site which consists of 6 amino acids.

SEQ ID NO:12 is the amino acid sequence of the PTPL1 phosphatase.

SEQ ID NO:13 is the nucleotide sequence of the expressed sequence tag identified by GenBank accession number T32506.

Detailed Description of the Invention

The present invention in one aspect involves the cloning of a cDNA encoding a PARG GTPase activating protein. The sequence of the human gene is presented as SEQ ID NO:1, and the predicted amino acid sequence of this gene's protein product is presented as SEQ ID NO:2. Analysis of the sequence by comparison to nucleic acid and protein databases determined that PARG has several domains in addition to the GAP domain. These include a cysteine-rich domain located directly N-terminal of the GAP domain, a ZPH domain similar to the ZK669.1 gene product of *C. elegans* (Wilson et al., *Nature* 368: 32-38, 1994), and a PDZ domain binding site.

The GAP activity of PARG was determined as reported in Example 7 below. The GAP activity of this protein is strongest on Rho GTPase *in vitro*. GAP activities were also detected on Rac and Cdc42 *in vitro*. Because these activities on Rac and Cdc42 were observed at higher PARG concentrations than needed for Rho GAP activity, it is likely that Rho is the preferred *in vivo* target of PARG.

A cysteine-rich domain is located directly N-terminal of the GAP domain of PARG. This domain has been identified in various proteins including most PKC isoforms (which have two copies each of the domain), the protooncogene products Vav and Raf, diacylglycerol kinase and chimaerins (reviewed by Newton, *Curr. Biol.* 5: 973-976, 1995). The cysteine-rich domain has

been shown to bind Zn^{2+} (Ahmed *et al.*, *Biochem J.* 280: 233-241, 1991), and the domains found in PKCs and in chimaerins also bind phorbol esters and diacylglycerol (Ahmed *et al.*, 1991; Ono *et al.*, *Proc. Natl. Acad. Sci. USA* 86: 4868-4871, 1989). Generation of diacylglycerol or addition of phorbol ester increase the affinity of PKC molecules for membranes, and the resulting translocation of PKC from the cytosol to the plasma membrane is likely to involve interactions between the cysteine-rich domains and membrane phospholipids (Newton, 1995; Zhang *et al.*, *Cell* 81: 917-924, 1995). The cysteine-rich domain of PARG may mediate regulatable binding to the membrane and could possibly also be involved in regulation of the GAP activity. Thus, a function of the cysteine-rich domain of PARG may be analogous to a function of n(αl)-chimaerin, a Rac-specific GAP, which contains a copy of a homologous cysteine-rich domain; it has been shown that phospholipids and phorbol esters regulate the GAP activity of n(αl)-chimaerin (Ahmed *et al.*, *J. Biol. Chem.* 268: 10709-10712, 1993).

In the N-terminal part of PARG, a region of about 300 amino acid residues with similarity (27 % identity) to the gene product of the *C. elegans* gene ZK669.1a was identified, and denoted ZPH region. The overall domain structure of the ZK669. 1 a gene product is similar to PARG and it is possible that PARG is the human homolog of the *C. elegans* ZK669.1 a gene product. However, the RhoGAP domain and the cysteine-rich domain of the ZK669.1 a gene product is not significantly more similar to PARG (29 % identity within the RhoGAP domains, 24 % idenity within the cysteine-rich domains) compared to other human proteins containing these domains (24-31 % identity within the RhoGAP domains and 16-27 % identity within the cysteine-rich domains).

PDZ domains have been identified in a diverse set of proteins (Ponting and Phillips, *Trends Biochem. Sci.* 20: 102-103, 1995). These proteins seem to be involved in signal transduction, and many of them, if not all, are found in structures at the plasma membrane. The size of the PDZ domain of about 90 amino acid residues, and its appearance in signal transduction proteins suggested that it, like SH2 and SH3 domains, can mediate direct interactions with other molecules. We have shown that PARG binds specificially to PDZ4 of PTPL1 and that the binding-site for binding to PDZ 4 resides in the four most C-terminal amino acid residues of PARG. PDZ domains can bind strongly to a short peptide of only four amino acid residues, and the carboxy-group and the side chain of the C-terminal valine residue is important for binding. The crystal structure of the third PDZ domain of PSD-95 binding to a peptide (Doyle *et al.*, 1996; Morais Cabral *et al.*, 1996) confirms these results and shows that the

last four C-terminal amino acid residues of the peptide bind in a cleft of the domain with the C-terminal valine buried in a shallow pocket. Thus, the PDZ domain functions as a C-terminal peptide binding module. Because PDZ 4 binds to PARG, a complex between PTPL1, PARG, and Rho can be formed. Protein tyrosine kinases have been implicated to act upstream and downstream of Rho (Nobes and Hall, *J. Cell Sci.* 108:225-233, 1995; Ridley, *BioEssays* 16:321-327, 1994). Thus, PTPL1 can function as a negative regulator of kinases in the Rho signal pathway, and in complex with PARG, which inactivates Rho itself, it can be a powerful inhibitor of Rho signals.

The invention thus involves in one aspect PARG polypeptides, genes encoding those polypeptides, functional modifications and variants of the foregoing, useful fragments of the foregoing, as well as therapeutics relating thereto.

Homologs and alleles of the PARG nucleic acids of the invention can be identified by conventional techniques. Thus, an aspect of the invention is those nucleic acid sequences which code for PARG polypeptides and which hybridize to a nucleic acid molecule consisting of the coding region of SEQ ID NO:1, under stringent conditions. The term "stringent conditions" as used herein refers to parameters with which the art is familiar. Nucleic acid hybridization parameters may be found in references which compile such methods, e.g. *Molecular Cloning: A Laboratory Manual*, J. Sambrook, et al., eds., Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989, or *Current Protocols in Molecular Biology*, F.M. Ausubel, et al., eds., John Wiley & Sons, Inc., New York. More specifically, stringent conditions, as used herein, refers, for example, to hybridization at 65°C in hybridization buffer (3.5 x SSC, 0.02% Ficoll, 0.02% polyvinyl pyrolidone, 0.02% Bovine Serum Albumin, 2.5mM NaH_2PO_4(pH7), 0.5% SDS, 2mM EDTA). SSC is 0.15M sodium chloride/0.15M sodium citrate, pH7; SDS is sodium dodecyl sulphate; and EDTA is ethylenediaminetetracetic acid. After hybridization, the membrane upon which the DNA is transferred is washed at 2 x SSC at room temperature and then at 0.1 x SSC/0.1 x SDS at temperatures up to 65°C.

There are other conditions, reagents, and so forth which can used, which result in a similar degree of stringency. The skilled artisan will be familiar with such conditions, and thus they are not given here. It will be understood, however, that the skilled artisan will be able to manipulate the conditions in a manner to permit the clear identification of homologs and alleles of PARG nucleic acids of the invention. The skilled artisan also is familiar with the methodology for screening cells and libraries for expression of such molecules which then are

routinely isolated, followed by isolation of the pertinent nucleic acid molecule and sequencing.

In general homologs and alleles typically will share at least 40% nucleotide identity and/or at least 50% amino acid identity to SEQ ID NO:1 and SEQ ID NO:2, respectively, in some instances will share at least 50% nucleotide identity and/or at least 65% amino acid identity and in still other instances will share at least 60% nucleotide identity and/or at least 75% amino acid identity. Watson-Crick complements of the foregoing nucleic acids also are embraced by the invention.

In screening for PARG proteins, a Southern blot may be performed using the foregoing conditions, together with a radioactive probe. After washing the membrane to which the DNA is finally transferred, the membrane can be placed against X-ray film to detect the radioactive signal.

The invention also includes degenerate nucleic acids which include alternative codons to those present in the native materials. For example, serine residues are encoded by the codons TCA, AGT, TCC, TCG, TCT and AGC. Each of the six codons is equivalent for the purposes of encoding a serine residue. Thus, it will be apparent to one of ordinary skill in the art that any of the serine-encoding nucleotide triplets may be employed to direct the protein synthesis apparatus, *in vitro* or *in vivo*, to incorporate a serine residue into an elongating PARG polypeptide. Similarly, nucleotide sequence triplets which encode other amino acid residues include, but are not limited to,: CCA, CCC, CCG and CCT (proline codons); CGA, CGC, CGG, CGT, AGA and AGG (arginine codons); ACA, ACC, ACG and ACT (threonine codons); AAC and AAT (asparagine codons); and ATA, ATC and ATT (isoleucine codons). Other amino acid residues may be encoded similarly by multiple nucleotide sequences. Thus, the invention embraces degenerate nucleic acids that differ from the biologically isolated nucleic acids in codon sequence due to the degeneracy of the genetic code.

The invention also provides isolated unique fragments of SEQ ID NO:1 or complements of SEQ ID NO:1. A unique fragment is one that is a 'signature' for the larger nucleic acid. It, for example, is long enough to assure that its precise sequence is not found in molecules outside of the PARG nucleic acids defined above. Unique fragments can be used as probes in Southern blot assays to identify such nucleic acids, or can be used in amplification assays such as those employing PCR. As known to those skilled in the art, large probes such as 200 nucleotides or more are preferred for certain uses such as Southern blots, while smaller fragments will be preferred for uses such as PCR. Unique fragments also can be used to produce fusion proteins

for generating antibodies or determining binding of the polypeptide fragments, as demonstrated in the Examples, or for generating immunoassay components. Likewise, unique fragments can be employed to produce nonfused fragments of the PARG polypeptides, useful, for example, in the preparation of antibodies, in immunoassays, and as a competitive binding partner of the PTPL1 phosphatase and/or other polypeptides which bind to the PARG polypeptides, for example, in therapeutic applications. Unique fragments further can be used as antisense molecules to inhibit the expression of PARG nucleic acids and polypeptides, particularly for therapeutic purposes as described in greater detail below.

As will be recognized by those skilled in the art, the size of the unique fragment will depend upon its conservancy in the genetic code. Thus, some regions of SEQ ID NO:1 and its complement will require longer segments to be unique while others will require only short segments, typically between 12 and 32 nucleotides (e.g. 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31 and 32 bases long). Excluding nucleic acid molecules consisting completely of the nucleotide sequence of SEQ ID NO:3 (GenBank accession number T32345), SEQ ID NO:4 (GenBank accession number Z28937), SEQ ID NO:5 (GenBank accession number Z28520) or SEQ ID NO:13 (GenBank accession number T32506) which overlaps SEQ ID NO:1, virtually any segment of the region of SEQ ID NO:1 beginning at nucleotide 184 and ending at nucleotide 3966, or complements thereof, that is 18 or more nucleotides in length will be unique. A fragment which is completely composed of the sequence of SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5 or SEQ ID NO:13 is one which does not include any of the nucleotides unique to PARG. Preferred longer unique fragments include those which are at least 50, 100, 150, 200, 250, 300, or 500 nucleotide in length. Particularly preferred are those unique fragments drawn completely from the portion of SEQ ID NO:3 which is not overlapped by the sequence of SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5 or SEQ ID NO:13. Those skilled in the art are well versed in methods for selecting such sequences, typically on the basis of the ability of the unique fragment to selectively distinguish the sequence of interest from non-PARG nucleic acids. A comparison of the sequence of the fragment to those on known data bases typically is all that is necessary, although *in vitro* confirmatory hybridization and sequencing analysis may be performed.

As mentioned above, the invention embraces antisense oligonucleotides that selectively bind to a nucleic acid molecule encoding a PARG polypeptide, to decrease GTPase activation by PARG or phosphatase binding by PARG. This is desirable in virtually any medical condition

wherein a reduction in GTPase activating activity of PARG is desirable, including to reduce Rho family protein signal transduction, or wherein a reduction in phosphatase binding by PARG is desirable. Antisense molecules, in this manner, can be used to slow down or arrest the proliferation of cancer cells *in vivo*.

As used herein, the term "antisense oligonucleotide" or "antisense" describes an oligonucleotide that is an oligoribonucleotide, oligodeoxyribonucleotide, modified oligoribonucleotide, or modified oligodeoxyribonucleotide which hybridizes under physiological conditions to DNA comprising a particular gene or to an mRNA transcript of that gene and, thereby, inhibits the transcription of that gene and/or the translation of that mRNA. The antisense molecules are designed so as to interfere with transcription or translation of a target gene upon hybridization with the target gene or transcript. Those skilled in the art will recognize that the exact length of the antisense oligonucleotide and its degree of complementarity with its target will depend upon the specific target selected, including the sequence of the target and the particular bases which comprise that sequence. It is preferred that the antisense oligonucleotide be constructed and arranged so as to bind selectively with the target under physiological conditions, i.e., to hybridize substantially more to the target sequence than to any other sequence in the target cell under physiological conditions. Based upon SEQ ID NO:1, or upon allelic or homologous genomic and/or cDNA sequences, one of skill in the art can easily choose and synthesize any of a number of appropriate antisense molecules for use in accordance with the present invention. In order to be sufficiently selective and potent for inhibition, such antisense oligonucleotides should comprise at least 10 and, more preferably, at least 15 consecutive bases which are complementary to the target, although in certain cases modified oligonucleotides as short as 7 bases in length have been used successfully as antisense oligonucleotides (Wagner et al., *Nature Biotechnol.* 14:840-844, 1996). Most preferably, the antisense oligonucleotides comprise a complementary sequence of 20-30 bases. Although oligonucleotides may be chosen which are antisense to any region of the gene or mRNA transcripts, in preferred embodiments the antisense oligonucleotides correspond to N-terminal or 5' upstream sites such as translation initiation, transcription initiation or promoter sites. In addition, 3'-untranslated regions may be targeted. Targeting to mRNA splicing sites has also been used in the art but may be less preferred if alternative mRNA splicing occurs. In addition, the antisense is targeted, preferably, to sites in which mRNA secondary structure is not expected (see, e.g., Sainio et al., *Cell Mol. Neurobiol.* 14(5):439-457, 1994) and at which proteins are not expected to bind. Finally,

although, SEQ ID NO:1 discloses a cDNA sequence, one of ordinary skill in the art may easily derive the genomic DNA corresponding to the cDNA of SEQ ID NO:1. Thus, the present invention also provides for antisense oligonucleotides which are complementary to the genomic DNA corresponding to SEQ ID NO:1. Similarly, antisense to allelic or homologous cDNAs and genomic DNAs are enabled without undue experimentation.

In one set of embodiments, the antisense oligonucleotides of the invention may be composed of "natural" deoxyribonucleotides, ribonucleotides, or any combination thereof. That is, the 5' end of one native nucleotide and the 3' end of another native nucleotide may be covalently linked, as in natural systems, via a phosphodiester internucleoside linkage. These oligonucleotides may be prepared by art recognized methods which may be carried out manually or by an automated synthesizer. They also may be produced recombinantly by vectors.

In preferred embodiments, however, the antisense oligonucleotides of the invention also may include "modified" oligonucleotides. That is, the oligonucleotides may be modified in a number of ways which do not prevent them from hybridizing to their target but which enhance their stability or targeting or which otherwise enhance their therapeutic effectiveness.

The term "modified oligonucleotide" as used herein describes an oligonucleotide in which (1) at least two of its nucleotides are covalently linked via a synthetic internucleoside linkage (i.e., a linkage other than a phosphodiester linkage between the 5' end of one nucleotide and the 3' end of another nucleotide) and/or (2) a chemical group not normally associated with nucleic acids has been covalently attached to the oligonucleotide. Preferred synthetic internucleoside linkages are phosphorothioates, alkylphosphonates, phosphorodithioates, phosphate esters, alkylphosphonothioates, phosphoramidates, carbamates, carbonates, phosphate triesters, acetamidates, carboxymethyl esters and peptides.

The term "modified oligonucleotide" also encompasses oligonucleotides with a covalently modified base and/or sugar. For example, modified oligonucleotides include oligonucleotides having backbone sugars which are covalently attached to low molecular weight organic groups other than a hydroxyl group at the 3' position and other than a phosphate group at the 5' position. Thus modified oligonucleotides may include a 2'-O-alkylated ribose group. In addition, modified oligonucleotides may include sugars such as arabinose instead of ribose. The present invention, thus, contemplates pharmaceutical preparations containing modified antisense molecules that are complementary to and hybridizable with, under physiological conditions, nucleic acids encoding PARG polypeptides, together with pharmaceutically acceptable carriers.

Antisense oligonucleotides may be administered as part of a pharmaceutical composition. Such a pharmaceutical composition may include the antisense oligonucleotides in combination with any standard physiologically and/or pharmaceutically acceptable carriers which are known in the art. The compositions should be sterile and contain a therapeutically effective amount of the antisense oligonucleotides in a unit of weight or volume suitable for administration to a patient. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredients. The term "physiologically acceptable" refers to a non-toxic material that is compatible with a biological system such as a cell, cell culture, tissue, or organism. The characteristics of the carrier will depend on the route of administration. Physiologically and pharmaceutically acceptable carriers include diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials which are well known in the art.

As used herein, a "vector" may be any of a number of nucleic acids into which a desired sequence may be inserted by restriction and ligation for transport between different genetic environments or for expression in a host cell. Vectors are typically composed of DNA although RNA vectors are also available. Vectors include, but are not limited to, plasmids, phagemids and virus genomes. A cloning vector is one which is able to replicate in a host cell, and which is further characterized by one or more endonuclease restriction sites at which the vector may be cut in a determinable fashion and into which a desired DNA sequence may be ligated such that the new recombinant vector retains its ability to replicate in the host cell. In the case of plasmids, replication of the desired sequence may occur many times as the plasmid increases in copy number within the host bacterium or just a single time per host before the host reproduces by mitosis. In the case of phage, replication may occur actively during a lytic phase or passively during a lysogenic phase. An expression vector is one into which a desired DNA sequence may be inserted by restriction and ligation such that it is operably joined to regulatory sequences and may be expressed as an RNA transcript. Vectors may further contain one or more marker sequences suitable for use in the identification of cells which have or have not been transformed or transfected with the vector. Markers include, for example, genes encoding proteins which increase or decrease either resistance or sensitivity to antibiotics or other compounds, genes which encode enzymes whose activities are detectable by standard assays known in the art (e.g., ß-galactosidase or alkaline phosphatase), and genes which visibly affect the phenotype of transformed or transfected cells, hosts, colonies or plaques (e.g., green fluorescent protein).

Preferred vectors are those capable of autonomous replication and expression of the structural gene products present in the DNA segments to which they are operably joined.

As used herein, a coding sequence and regulatory sequences are said to be "operably" joined when they are covalently linked in such a way as to place the expression or transcription of the coding sequence under the influence or control of the regulatory sequences. If it is desired that the coding sequences be translated into a functional protein, two DNA sequences are said to be operably joined if induction of a promoter in the 5' regulatory sequences results in the transcription of the coding sequence and if the nature of the linkage between the two DNA sequences does not (1) result in the introduction of a frame-shift mutation, (2) interfere with the ability of the promoter region to direct the transcription of the coding sequences, or (3) interfere with the ability of the corresponding RNA transcript to be translated into a protein. Thus, a promoter region would be operably joined to a coding sequence if the promoter region were capable of effecting transcription of that DNA sequence such that the resulting transcript might be translated into the desired protein or polypeptide.

The precise nature of the regulatory sequences needed for gene expression may vary between species or cell types, but shall in general include, as necessary, 5' non-transcribed and 5' non-translated sequences involved with the initiation of transcription and translation respectively, such as a TATA box, capping sequence, CAAT sequence, and the like. Especially, such 5' non-transcribed regulatory sequences will include a promoter region which includes a promoter sequence for transcriptional control of the operably joined gene. Regulatory sequences may also include enhancer sequences or upstream activator sequences as desired. The vectors of the invention may optionally include 5' leader or signal sequences. The choice and design of an appropriate vector is within the ability and discretion of one of ordinary skill in the art.

Expression vectors containing all the necessary elements for expression are commercially available and known to those skilled in the art. See, e.g., Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press, 1989. Cells are genetically engineered by the introduction into the cells of heterologous DNA (RNA) encoding PARG polypeptide or fragment or variant thereof. That heterologous DNA (RNA) is placed under operable control of transcriptional elements to permit the expression of the heterologous DNA in the host cell.

Preferred systems for mRNA expression in mammalian cells are those such as pRc/CMV (available from Invitrogen, Carlsbad, CA) that contain a selectable marker such as a gene that

confers G418 resistance (which facilitates the selection of stably transfected cell lines) and the human cytomegalovirus (CMV) enhancer-promoter sequences. Additionally, suitable for expression in primate or canine cell lines is the pCEP4 vector (Invitrogen), which contains an Epstein Barr virus (EBV) origin of replication, facilitating the maintenance of plasmid as a multicopy extrachromosomal element. Another expression vector is the pEF-BOS plasmid containing the promoter of polypeptide Elongation Factor 1α, which stimulates efficiently transcription *in vitro*. The plasmid is described by Mishizuma and Nagata (*Nuc. Acids Res.* 18:5322, 1990), and its use in transfection experiments is disclosed by, for example, Demoulin (*Mol. Cell. Biol.* 16:4710-4716, 1996). Still another preferred expression vector is an adenovirus, described by Stratford-Perricaudet, which is defective for E1 and E3 proteins (*J. Clin. Invest.* 90:626-630, 1992). The use of the adenovirus as an Adeno.P1A recombinant is disclosed by Warnier et al., in intradermal injection in mice for immunization against P1A (*Int. J. Cancer*, 67:303-310, 1996).

The invention also embraces so-called expression kits, which allow the artisan to prepare a desired expression vector or vectors. Such expression kits include at least separate portions of each of the previously discussed coding sequences. Other components may be added, as desired, as long as the previously mentioned sequences, which are required, are included.

The invention also permits the construction of PARG gene "knock-outs" in cells and in animals, providing materials for studying certain aspects of GTPase activating activity and signal transduction.

The invention also provides isolated polypeptides, which include the polypeptide of SEQ ID NO:2 and unique fragments of SEQ ID NO:2, particularly amino acids 193-509, 613-652 and 658-898 of SEQ ID NO:2, as well as the carboxyl terminal 4, 5 or 6 amino acids of SEQ ID NO:2. Such polypeptides are useful, for example, alone or as fusion proteins to generate antibodies, as a components of an immunoassay.

A unique fragment of an PARG polypeptide, in general, has the features and characteristics of unique fragments as discussed above in connection with nucleic acids. As will be recognized by those skilled in the art, the size of the unique fragment will depend upon factors such as whether the fragment constitutes a portion of a conserved protein domain. Thus, some regions of amino acids 658-898 of SEQ ID NO:2, amino acid residues 613-652 of SEQ ID NO:2 and amino acid residues of 193-509 SEQ ID NO:2, will require longer segments to be unique while others will require only short segments, typically between 5 and 12 amino acids (e.g. 5, 6,

7, 8, 9, 10, 11 and12 amino acids long). Virtually any segment of amino acids 658-898 of SEQ ID NO:2, amino acid residues 613-652 of SEQ ID NO:2 and amino acid residues of 193-509 SEQ ID NO:2, that is 10 or more amino acids in length will be unique.

Unique fragments of a polypeptide preferably are those fragments which retain a distinct functional capability of the polypeptide. Functional capabilities which can be retained in a unique fragment of a polypeptide include interaction with antibodies, interaction with other polypeptides (such as Rho) or fragments thereof, selective binding of nucleic acids or proteins (such as PTPL1), and enzymatic activity. Those skilled in the art are well versed in methods for selecting unique amino acid sequences, typically on the basis of the ability of the unique fragment to selectively distinguish the sequence of interest from non-family members. A comparison of the sequence of the fragment to those on known data bases typically is all that is necessary.

The invention embraces variants of the PARG polypeptides described above. As used herein, a "variant" of a PARG polypeptide is a polypeptide which contains one or more modifications to the primary amino acid sequence of a PARG polypeptide. Modifications which create a PARG variant can be made to a PARG polypeptide 1) to reduce or eliminate an activity of a PARG polypeptide, such as PTPL1 binding or GAP activity for Rho GTPase; 2) to enhance a property of a PARG polypeptide, such as protein stability in an expression system or the stability of protein-protein binding; or 3) to provide a novel activity or property to a PARG polypeptide, such as addition of an antigenic epitope or addition of a detectable moiety. Modifications to a PARG polypeptide are typically made to the nucleic acid which encodes the PARG polypeptide, and can include deletions, point mutations, truncations, amino acid substitutions and additions of amino acids or non-amino acid moieties. Alternatively, modifications can be made directly to the polypeptide, such as by cleavage, addition of a linker molecule, addition of a detectable moiety, such as biotin, addition of a fatty acid, and the like. Modifications also embrace fusion proteins comprising all or part of the PARG amino acid sequence.

In general, variants include PARG polypeptides which are modified specifically to alter a feature of the polypeptide unrelated to its physiological activity. For example, cysteine residues can be substituted or deleted to prevent unwanted disulfide linkages. Similarly, certain amino acids can be changed to enhance expression of a PARG polypeptide by eliminating proteolysis by proteases in an expression system (e.g., dibasic amino acid residues in yeast expression

systems in which KEX2 protease activity is present).

Mutations of a nucleic acid which encode a PARG polypeptide preferably preserve the amino acid reading frame of the coding sequence, and preferably do not create regions in the nucleic acid which are likely to hybridize to form secondary structures, such a hairpins or loops, which can be deleterious to expression of the variant polypeptide.

Mutations can be made by selecting an amino acid substitution, or by random mutagenesis of a selected site in a nucleic acid which encodes the polypeptide. Variant polypeptides are then expressed and tested for one or more activities to determine which mutation provides a variant polypeptide with the desired properties. Further mutations can be made to variants (or to non-variant PARG polypeptides) which are silent as to the amino acid sequence of the polypeptide, but which provide preferred codons for translation in a particular host. The preferred codons for translation of a nucleic acid in, e.g., *E. coli*, are well known to those of ordinary skill in the art. Still other mutations can be made to the noncoding sequences of a PARG gene or cDNA clone to enhance expression of the polypeptide. The activity of variants of PARG polypeptides can be tested by cloning the gene encoding the variant PARG polypeptide into a bacterial or mammalian expression vector, introducing the vector into an appropriate host cell, expressing the variant PARG polypeptide, and testing for a functional capability of the PARG polypeptides as disclosed herein. For example, the variant PARG polypeptide can be tested for Rho GAP activity as disclosed in Example 7, or for PDZ binding as disclosed in other Examples herein. Preparation of other variant polypeptides may favor testing of other activities, as will be known to one of ordinary skill in the art.

The skilled artisan will also realize that conservative amino acid substitutions may be made in PARG polypeptides to provide functionally equivalent variants of the foregoing polypeptides, i.e, the variants retain the functional capabilities of the PARG polypeptides. As used herein, a "conservative amino acid substitution" refers to an amino acid substitution which does not alter the relative charge or size characteristics of the protein in which the amino acid substitution is made. Variants can be prepared according to methods for altering polypeptide sequence known to one of ordinary skill in the art such as are found in references which compile such methods, e.g. *Molecular Cloning: A Laboratory Manual*, J. Sambrook, et al., eds., Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989, or *Current Protocols in Molecular Biology*, F.M. Ausubel, et al., eds., John Wiley & Sons, Inc., New York. Exemplary functionally equivalent variants of the PARG polypeptides include conservative

amino acid substitutions of SEQ ID NO:2, particularly conservative substitutions of amino acids other than 193-509, 613-652 or 658-898 of SEQ ID NO:2. However, conservative substitutions of amino acids 193-509, 613-652 or 658-898 of SEQ ID NO:2 can be made as well. Conservative substitutions of amino acids include substitutions made amongst amino acids within the following groups: (a) M, I, L, V; (b) F, Y, W; (c) K, R, H; (d) A, G; (e) S, T; (f) Q, N; and (g) E, D. Changes to the carboxyl terminal valine of the PARG PDZ domain binding site are not preferred for retention of maximal binding activity.

Conservative amino-acid substitutions in the amino acid sequence of PARG polypeptides to produce functionally equivalent variants of PARG polypeptides typically are made by alteration of the nucleic acid encoding PARG polypeptides (SEQ ID NO:1). Such substitutions can be made by a variety of methods known to one of ordinary skill in the art. For example, amino acid substitutions may be made by PCR-directed mutation, site-directed mutagenesis according to the method of Kunkel (Kunkel, *Proc. Nat. Acad. Sci. U.S.A.* 82: 488-492, 1985), or by chemical synthesis of a gene encoding a PARG polypeptide. Where amino acid substitutions are made to a small unique fragment of a PARG polypeptide, such as a PDZ-domain binding site peptide, the substitutions can be made by directly synthesizing the peptide. The activity of functionally equivalent fragments of PARG polypeptides can be tested by cloning the gene encoding the altered PARG polypeptide into a bacterial or mammalian expression vector, introducing the vector into an appropriate host cell, expressing the altered PARG polypeptide, and testing for a functional capability of the PARG polypeptides as disclosed herein. Peptides which are chemically synthesized can be tested directly for function, e.g., for binding to a PDZ 4 domain of PTPL1.

The invention as described herein has a number of uses, some of which are described elsewhere herein. First, the invention permits isolation of the PARG protein molecule (SEQ ID NO:2). A variety of methodologies well-known to the skilled practitioner can be utilized to obtain isolated PARG molecules. The polypeptide may be purified from cells which naturally produce the polypeptide by chromatographic means or immunological recognition. Alternatively, an expression vector may be introduced into cells to cause production of the polypeptide. In another method, mRNA transcripts may be microinjected or otherwise introduced into cells to cause production of the encoded polypeptide. Translation of mRNA in cell-free extracts such as the reticulocyte lysate system also may be used to produce polypeptide. Those skilled in the art also can readily follow known methods for isolating PARG polypeptides.

These include, but are not limited to, immunochromotography, HPLC, size-exclusion chromatography, ion-exchange chromatography and immune-affinity chromatography.

The isolation of the PARG gene also makes it possible for the artisan to diagnose a disorder characterized by expression of PARG. These methods involve determining expression of the PARG gene, and/or PARG polypeptides derived therefrom. In the former situation, such determinations can be carried out via any standard nucleic acid determination assay, including the polymerase chain reaction as exemplified in the examples below, or assaying with labeled hybridization probes.

The invention also makes it possible isolate proteins having a PDZ4 domain by the binding of such proteins to the PDZ domain binding site disclosed herein. The identification of the PDZ domain binding site also permits one of skill in the art to block the binding of a protein having a PDZ4 domain, such as PTPL1, with a binding partner having a PDZ4 domain binding site, such as PARG. Binding of the proteins can be effected by introducing into a biological system in which the proteins bind (e.g., a cell) a polypeptide including a PDZ domain binding site in an amount sufficient to block the binding. The identification of the PDZ4 domain binding site in PARG also enables one of skill in the art to prepare modified proteins, using standard recombinant DNA techniques, which can bind to proteins containing a PDZ4 domain. For example, when one desires to target a certain protein to the inner membrane surface where proteins containing a PDZ domain, such as PTPL1, are localized, one can prepare a fusion polypeptide of the protein and the PDZ4 domain binding site. Preferably, the PDZ domain binding site is fused to the carboxy terminus of the protein. Additional uses are described further herein.

The invention further provides methods for reducing or increasing Rho family signal transduction in a cell. Such methods are useful *in vitro* for altering the Rho signal transduction, for example, in testing compounds for potential to block aberrant Rho signal transduction. *In vivo*, such methods are useful for modulating actin polymerization, cell proliferation and release of secretory granules from mast cells (see, e.g., Price et al., *Curr. Biol.* 5:68-73, 1995), e.g., to treat allergy. Increasing Rho signal transduction in a cell by, e.g., introducing a dominant negative PARG polypeptide in the cell, can be used to provide a model system for testing the effects of putative inhibitors of Rho signal transduction. Such methods also are useful in the treatment of conditions which result from excessive or deficient Rho signal transduction. Rho signal transduction can be measured by studying actin reorganization or by measuring the ratio

of Rho-bound GTP/GDP. Various modulators of PARG GTPase activating activity can be screened for effects on Rho signal transduction using the methods disclosed herein. The skilled artisan can first determine the modulation of a PARG activity, such as GTPase activating activity, and then apply such a modulator to a target cell or subject and assess the effect on the target cell or subject. For example, in screeing for modulators of PARG useful in the treatment of mast cell secretion, mast cells in culture can be contacted with PARG modulators and the increase or decrease of secretory granule release by the mast cells can be determined according to standard procedures. PARG activity modulators can be assessed for their effects on other Rho signal transduction downstream effects by similar methods in other cell types.

The invention also provides, in certain embodiments, "dominant negative" polypeptides derived from SEQ ID NO:2. A dominant negative polypeptide is an inactive variant of a protein, which, by interacting with the cellular machinery, displaces an active protein from its interaction with the cellular machinery or competes with the active protein, thereby reducing the effect of the active protein. For example, a dominant negative receptor which binds a ligand but does not transmit a signal in response to binding of the ligand can reduce the biological effect of expression of the ligand. Likewise, a dominant negative catalytically-inactive kinase which interacts normally with target proteins but does not phosphorylate the target proteins can reduce phosphorylation of the target proteins in response to a cellular signal. Similarly, a dominant negative transcription factor which binds to a promoter site in the control region of a gene but does not increase gene transcription can reduce the effect of a normal transcription factor by occupying promoter binding sites without increasing transcription.

The end result of the expression of a dominant negative polypeptide in a cell is a reduction in function of active proteins. One of ordinary skill in the art can assess the potential for a dominant negative variant of a protein, and using standard mutagenesis techniques to create one or more dominant negative variant polypeptides. For example, given the teachings contained herein of a PARG polypeptide, one of ordinary skill in the art can modify the sequence of the PARG polypeptide by site-specific mutagenesis, scanning mutagenesis, partial gene deletion or truncation, and the like. See, e.g., U.S. Patent No. 5,580,723 and Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press, 1989. The skilled artisan then can test the population of mutagenized polypeptides for diminution in a selected activity (e.g., PARG GAP activity) and for retention of a desired activity (e.g., PARG binding to PTPL1). Other similar methods for creating and testing dominant negative variants of

a protein will be apparent to one of ordinary skill in the art.

Dominant negative PARG proteins include variants in which a portion of the PDZ4 domain binding site has been mutated or deleted to reduce or eliminate PARG interaction with PTPL1. Other examples include partial deletion PARG variants which have the GAP domain deleted. Such variants retain the capability to bind PTPL1 but cannot enhance GTPase activity in Rho. A GAP-negative PARG variant does not, therefore, stimulate downstream signal transduction pathways such as the Rho pathway.

The invention also involves agents such as polypeptides which bind to PARG polypeptides and to complexes of PARG polypeptides and their phosphatase binding partners. Such binding agents can be used, for example, in screening assays to detect the presence or absence of PARG polypeptides and complexes of PARG polypeptides and their phosphatase binding partners and in purification protocols to isolate PARG polypeptides and complexes of PARG polypeptides and their phosphatase binding partners. Such agents also can be used to inhibit the native activity of the PARG polypeptides or their phosphatase binding partners, for example, by binding to such polypeptides, or their binding partners or both.

The invention, therefore, embraces peptide binding agents which, for example, can be antibodies or fragments of antibodies having the ability to selectively bind to PARG polypeptides. Antibodies include polyclonal and monoclonal antibodies, prepared according to conventional methodology.

Significantly, as is well-known in the art, only a small portion of an antibody molecule, the paratope, is involved in the binding of the antibody to its epitope (see, in general, Clark, W.R. (1986) The Experimental Foundations of Modern Immunology Wiley & Sons, Inc., New York; Roitt, I. (1991) Essential Immunology, 7th Ed., Blackwell Scientific Publications, Oxford). The pFc' and Fc regions, for example, are effectors of the complement cascade but are not involved in antigen binding. An antibody from which the pFc' region has been enzymatically cleaved, or which has been produced without the pFc' region, designated an $F(ab')_2$ fragment, retains both of the antigen binding sites of an intact antibody. Similarly, an antibody from which the Fc region has been enzymatically cleaved, or which has been produced without the Fc region, designated an Fab fragment, retains one of the antigen binding sites of an intact antibody molecule. Proceeding further, Fab fragments consist of a covalently bound antibody light chain and a portion of the antibody heavy chain denoted Fd. The Fd fragments are the major

determinant of antibody specificity (a single Fd fragment may be associated with up to ten different light chains without altering antibody specificity) and Fd fragments retain epitope-binding ability in isolation.

Within the antigen-binding portion of an antibody, as is well-known in the art, there are complementarity determining regions (CDRs), which directly interact with the epitope of the antigen, and framework regions (FRs), which maintain the tertiary structure of the paratope (see, in general, Clark, 1986; Roitt, 1991). In both the heavy chain Fd fragment and the light chain of IgG immunoglobulins, there are four framework regions (FR1 through FR4) separated respectively by three complementarity determining regions (CDR1 through CDR3). The CDRs, and in particular the CDR3 regions, and more particularly the heavy chain CDR3, are largely responsible for antibody specificity.

It is now well-established in the art that the non-CDR regions of a mammalian antibody may be replaced with similar regions of conspecific or heterospecific antibodies while retaining the epitopic specificity of the original antibody. This is most clearly manifested in the development and use of "humanized" antibodies in which non-human CDRs are covalently joined to human FR and/or Fc/pFc' regions to produce a functional antibody. Thus, for example, PCT International Publication Number WO 92/04381 teaches the production and use of humanized murine RSV antibodies in which at least a portion of the murine FR regions have been replaced by FR regions of human origin. Such antibodies, including fragments of intact antibodies with antigen-binding ability, are often referred to as "chimeric" antibodies.

Thus, as will be apparent to one of ordinary skill in the art, the present invention also provides for $F(ab')_2$, Fab, Fv and Fd fragments; chimeric antibodies in which the Fc and/or FR and/or CDR1 and/or CDR2 and/or light chain CDR3 regions have been replaced by homologous human or non-human sequences; chimeric $F(ab')_2$ fragment antibodies in which the FR and/or CDR1 and/or CDR2 and/or light chain CDR3 regions have been replaced by homologous human or non-human sequences; chimeric Fab fragment antibodies in which the FR and/or CDR1 and/or CDR2 and/or light chain CDR3 regions have been replaced by homologous human or non-human sequences; and chimeric Fd fragment antibodies in which the FR and/or CDR1 and/or CDR2 regions have been replaced by homologous human or non-human sequences. The present invention also includes so-called single chain antibodies.

Thus, the invention involves polypeptides of numerous size and type that bind specifically to PARG polypeptides, and complexes of both PARG polypeptides and their

phosphatase binding partners. These polypeptides may be derived also from sources other than antibody technology. For example, such polypeptide binding agents can be provided by degenerate peptide libraries which can be readily prepared in solution, in immobilized form or as phage display libraries. Combinatorial libraries also can be synthesized of peptides containing one or more amino acids. Libraries further can be synthesized of peptoids and non-peptide synthetic moieties.

Phage display can be particularly effective in identifying binding peptides useful according to the invention. Briefly, one prepares a phage library (using e.g. m13, fd, or lambda phage), displaying inserts from 4 to about 80 amino acid residues using conventional procedures. The inserts may represent, for example, a completely degenerate or biased array. One then can select phage-bearing inserts which bind to the PARG polypeptide. This process can be repeated through several cycles of reselection of phage that bind to the PARG polypeptide. Repeated rounds lead to enrichment of phage bearing particular sequences. DNA sequence analysis can be conducted to identify the sequences of the expressed polypeptides. The minimal linear portion of the sequence that binds to the PARG polypeptide can be determined. One can repeat the procedure using a biased library containing inserts containing part or all of the minimal linear portion plus one or more additional degenerate residues upstream or downstream thereof. Yeast two-hybrid screening methods also may be used to identify polypeptides that bind to the PARG polypeptides. Thus, the PARG polypeptides of the invention, or a fragment thereof, can be used to screen peptide libraries, including phage display libraries, to identify and select peptide binding partners of the PARG polypeptides of the invention. Such molecules can be used, as described, for screening assays, for purification protocols, for interfering directly with the functioning of PARG and for other purposes that will be apparent to those of ordinary skill in the art.

A PARG polypeptide, or a fragment which contains the C-terminal PDZ4 domain binding site, also can be used to isolate their native binding partners, including, e.g., the PTPL1 phosphatase that complexes with PARG. Isolation of phosphatases may be performed according to well-known methods. For example, isolated PARG polypeptides can be attached to a substrate, and then a solution suspected of containing the phosphatase may be applied to the substrate. If the phosphatase binding partner for PARG polypeptides is present in the solution, then it will bind to the substrate-bound PARG polypeptide. The phosphatase then may be isolated. Other proteins which are binding partners for PARG, such as other proteins which

contain PDZ4 domains may be isolated by similar methods without undue experimentation. Similarly, other proteins which bind PARG (e.g. Rho) can be isolated from biological samples and/or extracts by such methods.

Isolation of the PARG protein enables the skilled artisan to use the protein for isolation of molecules which bind to it. For example, isolated PARG can be used to isolate PTPL1 and other proteins which contain PDZ4 domains. The PARG or PDZ binding fragment can be immobilized on chromatographic media, such as polystyrene beads, or a filter, and the immobilized protein can be used to isolate proteins containing a PDZ4 domain from biological samples with no more than routine experimentation according to art-standard procedures for affinity chromatography. Such procedures are described in greater detail below.

It will also be recognized that the invention embraces the use of the PARG cDNA sequences in expression vectors, as well as to transfect host cells and cell lines, be these prokaryotic (e.g., *E. coli*), or eukaryotic (e.g., CHO cells, COS cells, yeast expression systems and recombinant baculovirus expression in insect cells). Especially useful are mammalian cells such as mouse, hamster, pig, goat, primate, etc. They may be of a wide variety of tissue types, and include primary cells and cell lines. Specific examples include dendritic cells, U293 cells, peripheral blood leukocytes, bone marrow stem cells and embryonic stem cells. The expression vectors require that the pertinent sequence, i.e., those nucleic acids described *supra*, be operably linked to a promoter.

When administered, the therapeutic compositions of the present invention are administered in pharmaceutically acceptable preparations. Such preparations may routinely contain pharmaceutically acceptable concentrations of salt, buffering agents, preservatives, compatible carriers, supplementary immune potentiating agents such as adjuvants and cytokines and optionally other therapeutic agents.

The therapeutics of the invention can be administered by any conventional route, including injection or by gradual infusion over time. The administration may, for example, be oral, intravenous, intraperitoneal, intramuscular, intracavity, subcutaneous, or transdermal. When antibodies are used therapeutically, a preferred route of administration is by pulmonary aerosol. Techniques for preparing aerosol delivery systems containing antibodies are well known to those of skill in the art. Generally, such systems should utilize components which will not significantly impair the biological properties of the antibodies, such as the paratope binding capacity (see, for example, Sciarra and Cutie, "Aerosols," in Remington's Pharmaceutical

Sciences, 18th edition, 1990, pp 1694-1712; incorporated by reference). Those of skill in the art can readily determine the various parameters and conditions for producing antibody aerosols without resort to undue experimentation. When using antisense preparations of the invention, slow intravenous administration is preferred.

Preparations for parenteral administration include sterile aqueous or non-aqueous solutions, suspensions, and emulsions. Examples of non-aqueous solvents are propylene glycol, polyethylene glycol, vegetable oils such as olive oil, and injectable organic esters such as ethyl oleate. Aqueous carriers include water, alcoholic/aqueous solutions, emulsions or suspensions, including saline and buffered media. Parenteral vehicles include sodium chloride solution, Ringer's dextrose, dextrose and sodium chloride, lactated Ringer's or fixed oils. Intravenous vehicles include fluid and nutrient replenishers, electrolyte replenishers (such as those based on Ringer's dextrose), and the like. Preservatives and other additives may also be present such as, for example, antimicrobials, anti-oxidants, chelating agents, and inert gases and the like.

The preparations of the invention are administered in effective amounts. An effective amount is that amount of a pharmaceutical preparation that alone, or together with further doses, produces the desired response. In the case of treating cancer, the desired response is inhibiting the progression of the cancer. This may involve only slowing the progression of the disease temporarily, although more preferably, it involves halting the progression of the disease permanently. This can be monitored by routine methods or can be monitored according to diagnostic methods of the invention discussed herein. Other therapeutic uses of PARG include the modulation of actin reorganization, and modulation of mast cell secretory granule release to treat allergic responses.

The invention also contemplates gene therapy. The procedure for performing *ex vivo* gene therapy is outlined in U.S. Patent 5,399,346 and in exhibits submitted in the file history of that patent, all of which are publicly available documents. In general, it involves introduction *in vitro* of a functional copy of a gene into a cell(s) of a subject which contains a defective copy of the gene, and returning the genetically engineered cell(s) to the subject. The functional copy of the gene is under operable control of regulatory elements which permit expression of the gene in the genetically engineered cell(s). Numerous transfection and transduction techniques as well as appropriate expression vectors are well known to those of ordinary skill in the art, some of which are described in PCT application WO95/00654. *In vivo* gene therapy using vectors such as adenovirus, retroviruses, herpes virus, and targeted liposomes also is contemplated according to

the invention.

The invention further provides efficient methods of identifying pharmacological agents or lead compounds for agents active at the level of a PARG or PARG fragment modulatable cellular function. In particular, such functions include Rho signal transduction and formation of a PTPL1-PARG protein complex. Generally, the screening methods involve assaying for compounds which interfere with a PARG activity such as PARG-PTPL1 binding, etc. Such methods are adaptable to automated, high throughput screening of compounds. The target therapeutic indications for pharmacological agents detected by the screening methods are limited only in that the target cellular function be subject to modulation by alteration of the formation of a complex comprising a PARG polypeptide or fragment thereof and one or more natural PARG intracellular binding targets, such as PTPL1 or other protein including a PDZ 4 domain. Target indications include cellular processes modulated by Rho signal transduction following receptor-ligand binding and PTPL1-mediated phosphorylation.

A wide variety of assays for pharmacological agents are provided, including, labeled *in vitro* protein-protein binding assays, electrophoretic mobility shift assays, immunoassays, cell-based assays such as two- or three-hybrid screens, expression assays, etc. For example, three-hybrid screens are used to rapidly examine the effect of transfected nucleic acids on the intracellular binding of PARG or PARG fragments to specific intracellular targets. The transfected nucleic acids can encode, for example, combinatorial peptide libraries or antisense molecules. Convenient reagents for such assays, e.g., GAL4 fusion proteins, are known in the art. An exemplary cell-based assay involves transfecting a cell with a nucleic acid encoding a PTPL1-binding PARG polypeptide (e.g., including a PDZ domain binding site) fused to a GAL4 DNA binding domain and a nucleic acid encoding a PTPL1 PDZ 4 domain fused to a transcription activation domain such as VP16. The cell also contains a reporter gene operably linked to a gene expression regulatory region, such as one or more GAL4 binding sites. Activation of reporter gene transcription occurs when the PARG and PTPL1 PDZ 4 fusion polypeptides bind such that the GAL4 DNA binding domain and the VP16 transcriptional activation domain are brought into proximity to enable transcription of the reporter gene. Agents which modulate a PARG polypeptide mediated cell function are then detected through a change in the expression of reporter gene. Methods for determining changes in the expression of a reporter gene are known in the art.

PARG fragments used in the methods, when not produced by a transfected nucleic acid

are added to an assay mixture as an isolated polypeptide. PARG polypeptides preferably are produced recombinantly, although such polypeptides may be isolated from biological extracts. Recombinantly produced PARG polypeptides include chimeric proteins comprising a fusion of a PARG protein with another polypeptide, e.g., a polypeptide capable of providing or enhancing protein-protein binding, sequence specific nucleic acid binding (such as GAL4), enhancing stability of the PARG polypeptide under assay conditions, or providing a detectable moiety, such as green fluorescent protein. A polypeptide fused to a PARG polypeptide or fragment may also provide means of readily detecting the fusion protein, e.g., by immunological recognition or by fluorescent labeling.

The assay mixture is comprised of a natural intracellular PARG binding target such as a Rho protein, PTPL1 protein or fragment thereof capable of binding to PARG. While natural PARG binding targets may be used, it is frequently preferred to use portions (e.g., peptides or nucleic acid fragments) or analogs (i.e., agents which mimic the PARG binding properties of the natural binding target for purposes of the assay) of the PARG binding target so long as the portion or analog provides binding affinity and avidity to the PARG fragment measurable in the assay.

The assay mixture also comprises a candidate pharmacological agent. Typically, a plurality of assay mixtures are run in parallel with different agent concentrations to obtain a different response to the various concentrations. Typically, one of these concentrations serves as a negative control, i.e., at zero concentration of agent or at a concentration of agent below the limits of assay detection. Candidate agents encompass numerous chemical classes, although typically they are organic compounds. Preferably, the candidate pharmacological agents are small organic compounds, i.e., those having a molecular weight of more than 50 yet less than about 2500, preferably less than about 1000 and, more preferably, less than about 500. Candidate agents comprise functional chemical groups necessary for structural interactions with polypeptides and/or nucleic acids, and typically include at least an amine, carbonyl, hydroxyl or carboxyl group, preferably at least two of the functional chemical groups and more preferably at least three of the functional chemical groups. The candidate agents can comprise cyclic carbon or heterocyclic structure and/or aromatic or polyaromatic structures substituted with one or more of the above-identified functional groups. Candidate agents also can be biomolecules such as peptides, saccharides, fatty acids, sterols, isoprenoids, purines, pyrimidines, derivatives or structural analogs of the above, or combinations thereof and the like. Where the agent is a

nucleic acid, the agent typically is a DNA or RNA molecule, although modified nucleic acids as defined herein are also contemplated.

Candidate agents are obtained from a wide variety of sources including libraries of synthetic or natural compounds. For example, numerous means are available for random and directed synthesis of a wide variety of organic compounds and biomolecules, including expression of randomized oligonucleotides, synthetic organic combinatorial libraries, phage display libraries of random peptides, and the like. Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant and animal extracts are available or readily produced. Additionally, natural and synthetically produced libraries and compounds can be readily be modified through conventional chemical, physical, and biochemical means. Further, known pharmacological agents may be subjected to directed or random chemical modifications such as acylation, alkylation, esterification, amidification, etc. to produce structural analogs of the agents.

A variety of other reagents also can be included in the mixture. These include reagents such as salts, buffers, neutral proteins (e.g., albumin), detergents, etc. which may be used to facilitate optimal protein-protein and/or protein-nucleic acid binding. Such a reagent may also reduce non-specific or background interactions of the reaction components. Other reagents that improve the efficiency of the assay such as protease, inhibitors, nuclease inhibitors, antimicrobial agents, and the like may also be used.

The mixture of the foregoing assay materials is incubated under conditions whereby, but for the presence of the candidate pharmacological agent, the PARG polypeptide specifically binds the cellular binding target, a portion thereof or analog thereof. The order of addition of components, incubation temperature, time of incubation, and other perimeters of the assay may be readily determined. Such experimentation merely involves optimization of the assay parameters, not the fundamental composition of the assay. Incubation temperatures typically are between 4°C and 40°C. Incubation times preferably are minimized to facilitate rapid, high throughput screening, and typically are between 0.1 and 10 hours.

After incubation, the presence or absence of specific binding between the PARG polypeptide and one or more binding targets is detected by any convenient method available to the user. For cell free binding type assays, a separation step is often used to separate bound from unbound components. The separation step may be accomplished in a variety of ways. Conveniently, at least one of the components is immobilized on a solid substrate, from which the

unbound components may be easily separated. The solid substrate can be made of a wide variety of materials and in a wide variety of shapes, e.g., microtiter plate, microbead, dipstick, resin particle, etc. The substrate preferably is chosen to maximum signal to noise ratios, primarily to minimize background binding, as well as for ease of separation and cost.

Separation may be effected for example, by removing a bead or dipstick from a reservoir, emptying or diluting a reservoir such as a microtiter plate well, rinsing a bead, particle, chromotograpic column or filter with a wash solution or solvent. The separation step preferably includes multiple rinses or washes. For example, when the solid substrate is a microtiter plate, the wells may be washed several times with a washing solution, which typically includes those components of the incubation mixture that do not participate in specific bindings such as salts, buffer, detergent, non-specific protein, etc. Where the solid substrate is a magnetic bead, the beads may be washed one or more times with a washing solution and isolated using a magnet.

Detection may be effected in any convenient way for cell-based assays such as two- or three-hybrid screens. The transcript resulting from a reporter gene transcription assay of PARG polypeptide binding to a target molecule typically encodes a directly or indirectly detectable product, e.g., β-galactosidase activity, luciferase activity, and the like. For cell free binding assays, one of the components usually comprises, or is coupled to, a detectable label. A wide variety of labels can be used, such as those that provide direct detection (e.g., radioactivity, luminescence, optical or electron density, etc). or indirect detection (e.g., epitope tag such as the FLAG epitope, enzyme tag such as horseseradish peroxidase, etc.). The label may be bound to a PARG binding partner, or incorporated into the structure of the binding partner.

A variety of methods may be used to detect the label, depending on the nature of the label and other assay components. For example, the label may be detected while bound to the solid substrate or subsequent to separation from the solid substrate. Labels may be directly detected through optical or electron density, radioactive emissions, nonradiative energy transfers, etc. or indirectly detected with antibody conjugates, strepavidin-biotin conjugates, etc. Methods for detecting the labels are well known in the art.

The invention provides PARG-specific binding agents, methods of identifying and making such agents, and their use in diagnosis, therapy and pharmaceutical development. For example, PARG-specific pharmacological agents are useful in a variety of diagnostic and therapeutic applications, especially where disease or disease prognosis is associated with improper utilization of a pathway involving PARG, e.g., Rho activation, PTPL1-PARG complex

formation, etc. Novel PARG-specific binding agents include PARG-specific antibodies and other natural intracellular binding agents identified with assays such as two hybrid screens, and non-natural intracellular binding agents identified in screens of chemical libraries and the like.

In general, the specificity of PARG binding to a binding agent is shown by binding equilibrium constants. Targets which are capable of selectively binding a PARG polypeptide preferably have binding equilibrium constants of at least about 10^7 M^{-1}, more preferably at least about 10^8 M^{-1}, and most preferably at least about 10^9 M^{-1}. The wide variety of cell based and cell free assays may be used to demonstrate PARG-specific binding. Cell based assays include one, two and three hybrid screens, assays in which PARG-mediated transcription is inhibited or increased, etc. Cell free assays include PARG-protein binding assays, immunoassays, etc. Other assays useful for screening agents which bind PARG polypeptides include fluorescence resonance energy transfer (FRET), and electrophoretic mobility shift analysis (EMSA).

Various techniques may be employed for introducing nucleic acids of the invention into cells, depending on whether the nucleic acids are introduced *in vitro* or *in vivo* in a host. Such techniques include transfection of nucleic acid-$CaPO_4$ precipitates, transfection of nucleic acids associated with DEAE, transfection with a retrovirus including the nucleic acid of interest, liposome mediated transfection, and the like. For certain uses, it is preferred to target the nucleic acid to particular cells. In such instances, a vehicle used for delivering a nucleic acid of the invention into a cell (e.g., a retrovirus, or other virus; a liposome) can have a targeting molecule attached thereto. For example, a molecule such as an antibody specific for a surface membrane protein on the target cell or a ligand for a receptor on the target cell can be bound to or incorporated within the nucleic acid delivery vehicle. For example, where liposomes are employed to deliver the nucleic acids of the invention, proteins which bind to a surface membrane protein associated with endocytosis may be incorporated into the liposome formulation for targeting and/or to facilitate uptake. Such proteins include capsid proteins or fragments thereof tropic for a particular cell type, antibodies for proteins which undergo internalization in cycling, proteins that target intracellular localization and enhance intracellular half life, and the like. Polymeric delivery systems also have been used successfully to deliver nucleic acids into cells, as is known by those skilled in the art. Such systems even permit oral delivery of nucleic acids.

Examples

Example 1: Production of PDZ Fusion Proteins

To identify proteins that bind to the PDZ domains of PTPL1, regions of PTPL1 cDNA corresponding to the various PDZ domains were produced by polymerase chain reaction and subcloned into the GST fusion protein expression vector pGEX1λT (Pharmacia): GST-PDZ 1. amino acid residues 1066-1166 of PTPL1; GST-PDZ 2-3. residues 1340-1579; GST-PDZ 3, residues 1469-1579; GST-PDZ 4, residues 1762-1864; GST-PDZ 4-5, residues 1762-1960 and GST-PDZ 5, residues 1856-1960 (Figure 1A). Domains and motifs indicated in Figure 1A are: L, leucine zipper motif; Band 4.1, a domain of 300 amino acid residues with homology to the Band 4.1 superfamily; P, PDZ domain; PTP, protein tyrosine phosphatase catalytic domain; GST, glutathione S-transferase. The different expression vector constructs were transformed into *E. coli*. Glutathione S-transferase (GST) fusion proteins were produced and purified as described by Ridley and Hall (*Cell* 70: 389-399, 1992) and then subjected to sodium dodecyl sulfate (SDS)-gel electrophoresis. Figure 1B shows that pure preparations of fusion proteins with expected sizes were obtained.

Example 2: Identification of Proteins Which Bind to PDZ4

PC-3 cells were obtained from American Type Culture Collection (Rockville, MD) and cultured as described (Saras *et al.*, 1994). Metabolic labeling of PC-3 cells was performed for 4 h in methionine- and cysteine-free MCDB 104 medium (Gibco/Life Technologies, Gaithersburg, MD) with 150 Ci/ml of ^{35}S-methionine and ^{35}S-cysteine (*in vivo* labeling mix; Amersham, Arlington Heights, IL). After labeling, the cells were solubilized in buffer containing 20 mM Tris-HCl, pH 7.4,150 mM NaCl, 10 mM EDTA, 0.5% Triton X-100, 0.5% deoxycholate, l mM dithiothreitol, 1.5% Trasylol (Bayer, Germany) and 1 mM phenylmethylsulfonyl fluoride (Sigma, St. Louis, MO). After 15 min on ice, cell debris was removed by centrifugation. Samples (1 ml) were then incubated for 1.5 h at 4°C with 10 μg of GST-PDZ fusion proteins bound to glutathione-Sepharose 4B beads (Pharmacia). The beads were pelleted and washed four times with solubilization buffer. The protein complexes were eluted by boiling for 5 min in SDS-sample buffer (100 mM Tris-HCl, pH 8.8, 0.01 % bromophenol blue, 36% glycerol, 4% SDS, 10 mM dithiothreitol) and analyzed by SDS-gel electrophoresis using 5-12 % polyacrylamide gels (Blobel and Dobberstein, *J. Cell Biol.* 67: 835-851, 1975). The gel was fixed, incubated with Amplify (Amersham) for 20 min, dried and subjected to fluorography. A

component of 150 kDa that bound to the fusion proteins GST-PDZ 4 and GST-PDZ-4-5 could be observed (Figure 2); this component did not bind to GST fusion proteins containing PDZ domains 1, 2, 3 or 5 only, thus indicating that the 150 kDa component interacts specifically with PDZ 4 of PTPL1.

Example 3: Purification of 150kDa Protein which binds to PDZ4

In order to characterize the 150 kDa component further, it was purified from PC-3 cells. Briefly, immobilized fusion protein GST-PDZ 4 was incubated with cell lysate from 1750 cm^2 of confluent PC-3 cells solubilized as described above. Samples (20 ml) were incubated for 1.5 h at 4°C with 200 μg of GST-PDZ 4 fusion protein bound to glutathione-Sepharose 4B beads. The beads were washed and the bound proteins were eluted and subjected to SDS-get electrophoresis as described above.

After staining of the gel with Coomassie Brilliant Blue, the band that contained the 150 kDa component was excised and subjected to in-gel digestion using modified trypsin or EndoLysC protease. The band containing the 150 kDa component was transferred to Eppendorf tubes and subjected to in-gel digestion (Hellman *et al.*, *Anal. Biochem.* 224: 451-455, 1995). In brief, the gel piece was washed with 0.2 M ammonium bicarbonate (for digestion with trypsin) or 0.5 M Tris-HCl pH 9.2 (for digestion with EndoLysC protease) and 50 % acetonitrile, then dried completely. During rehydration, 0.5 μg of modified trypsin, sequence grade (Promega, Madison, WI) or 0.5 μg of EndoLysC (WAKO Chemicals, Richmond, VA) was added and 0.2 M ammonium bicarbonate (for trypsin) or 0.1 M Tris-HCl pH 9.2 (for EndoLysC) was added in aliquots until the gel piece was immersed. After overnight incubation at 30°C, the supernatant was saved and combined with two further extractions from the gel piece. Generated peptides were isolated by reversed phase liquid chromatography using the SMART System (Pharmacia Biotech, Uppsala, Sweden). Peptides were sequenced on an Applied Biosystems (Foster City, CA) model 470A or 476A, following the manufacturers instructions.

Sequences were obtained from 10 peptides, and searches in different databases showed that none of these sequences were found in any known gene or protein, but the human Expressed Sequence Tags (ESTs) with GenBank accession numbers T32345, Z28937 and Z28520 (SEQ ID NOs:3, 4, 5), contained cDNA sequences corresponding to three of the obtained peptides. Oligonucleotides corresponding to the nucleotide sequences of the ESTs were designed and used as probes for Northern blots and screening of cDNA libraries.

Example 4: cDNA Cloning of PARG

The EST-derived oligonucleotides described above were used to screen different human cDNA libraries. Briefly, complementary and overlapping oligonucleotides corresponding to nucleotides 2-41 and 68-29 of an EST with the GenBank accession number Z28520 (SEQ ID NO:5) were made using a DNA synthesizer and labeled by a fill-in method (Sambrook *et al.*, 1989) using the Klenow fragment of DNA polymerase I (Amersham) and α-^{32}P-dCTP (3000Ci/mmol, Amersham). A λgt11 human skeletal muscle cDNA library (HL5002b; Clontech, Palo Alto, CA) was screened as described (Saras *et al.*, 1994), using the ^{32}P-labeled oligonucleotides as a probe. A positive clone was isolated, subcloned into pBluescript SK (Stratagene, La Jolla, CA) and thereafter sequenced.

Nucleotide sequencing revealed that the clone had a total length of 5237 bp with an open reading frame of 3783 bp, coding for a protein of 1261 amino acid residues. The open reading frame is flanked by a 5' untranslated sequence of 183 bp that contains an in frame stop codon at positions 166-168, and a 3' untranslated sequence of 1270 bp that has a poly(A) tail. The calculated molecular mass of the translated product is 142 kDa and the protein was, for reasons described below, denoted PARG. The amino acid sequence of PARG (SEQ ID NO:2) is shown in Figure 3A; the nucleotide sequence (SEQ ID NO:1) has been deposited in the EMBL database.

Example 5: Structure of the PARG Protein

The amino acid sequence of PARG contained all peptide sequences obtained previously (Figure 3A). In the deduced amino acid sequence of PARG no transmembrane domain or signal sequence for secretion were found, indicating that PARG is likely an intracellular protein. Three regions with homologies to other proteins could be identified: A GAP domain with similarity (23-33 % amino acid sequence identity) to proteins of the RhoGAP family (Lamarche and Hall, 1994) is found at amino acid residues 666-853, a cysteine-rich region at amino acid residues 613-652 has homology to a regulatory, phorbol ester-, diacylglycerol- and Zn2+- binding domain of members of the protein kinase C (PKC) family (Newton, 1995), and a region at amino acid residues 193-509 has homology (27 % identity) to the gene product of the *C. elegans* gene ZK669.1 a (EMBL accession number Z37093). Figure 3B shows an alignment of the latter homology region, denoted ZPH region(for ZK667.la-PARG homology). The alignment was done using the Clustal method (Higgins and Sharp,*CABIOS* 5: 151-153, 1989), with some

manual adjustment. Identical amino acid residues are boxed. Like PARG, the gene product of ZK669.1 a contains in addition to the ZPH region, a cysteine-rich domain and a GAP domain (Figure 3C). Domains and motifs indicated in Figure 3C are: ZPH, ZK669.1a-PARG Homology region; C, cysteine-rich domain; GAP, RhoGAP domain.

Example 6: Expression of PARG mRNA

Northern blot analysis was performed to determine expression of the PARG mRNA. A Northern blot filter with mRNA from different human tissues was purchased from Clontech. Each lane contained 2 μg of polyadenylated RNA from the indicated tissues. The filter was hybridized with the ^{32}P-labeled oligonucleotide probe described above, at 42°C overnight in a hybridization solution containing 50% formamide, 5 x SSC (lx SSC is 15 mM sodium citrate and 150 mM sodium chloride), 2 x Denhardt's solution, 0.5% SDS, 50 mM sodium phosphate, pH 6.9, and 0.1 mg/ml salmon sperm DNA. The filter was washed two times in 0.5 x SSC, 0.1% SDS at 55°C for 15 min. After washing, the filter was exposed to Amersham Hyperfilm MP.

Northern blot analysis of mRNA from various human tissues showed that a single PARG transcript of 5.5 kb was found in all screened tissues (Figure 4). The expression of PARG mRNA was high in skeletal muscle and heart and moderate in placenta, liver and pancreas. Low expression was observed in brain, lung and kidney. The size of the transcript suggested that the cDNA clone obtained was close to full length.

Example 7: GAP activity of PARG

In order to determine the GAP activity of PARG on proteins of the Rho family, the GAP domain of PARG was produced as a GST fusion protein in *E. coli* (Figure 5A). Briefly, a DNA fragment coding for the GAP domain, i.e., amino acid residues 658-898, of PARG was produced by polymerase chain reaction and subcloned into pGEX1λT and referred to as GST-GAP. pGEX2T-based expression vectors containing RhoA, Rac1 and Cdc42 (G25K isoform) cDNAs were obtained from Dr. A. Hall (MRC Laboratory for Molecular Cell Biology and Department of Biochemistry, University College London, UK). These different expression vector constructs were transformed into *E. coli*. The GST fusion proteins were produced and purified essentially as described above in Example 1. Recombinant Rho, Rac and Cdc42 proteins were subjected to thrombin cleavage (Ridley and Hall, 1992).

Recombinant Rho, Rac and Cdc42 were preloaded with γ-^{32}P-GTP and incubated for

various time periods in the presence of the GST-GAP fusion protein or, as control, GST protein. Thereafter, the radioactivity bound to the GTPase was determined as a measurement of the GTP hydrolysis activity. Briefly, 200 nM aliquots of recombinant Rho, Rac and Cdc42 were incubated at 30°C with 10 μCi γ-^{32}P-GTP in 20 mM Tris-HCl, pH 7.5, 25 mM NaCl, 4 mM EDTA, 0.1 mM dithiothreitol, and the nucleotide exchange was stopped after 10 min by the addition of 17 mM $MgCl_2$. Proteins (100 nM GST, 1 nM or 20 nM of GST-GAP fusion protein) were added to the reaction mixture and aliquots of 5 μl were withdrawn and collected on nitrocellulose filters (HA, Millipore, Bedford, MA) at 3 min intervals. The filters were washed with cold buffer (50 mM Tris-HCl pH 7.5, 50 mM NaCl, 5 mM $MgCl_2$), dried and subjected to scintillation counting. The amount of protein-bound radioactivity is expressed as the percentage of the total input.

The results show that the GAP domain of PARG, at the concentration of 1 nM, had a strong GAP activity on Rho (Figure 5B). At this concentration, no GAP activity on Rac or Cdc42 was detected (Figure 5C and 5D). However, at a concentration of 20 nM, the GST-GAP fusion protein was also active on Rac and Cdc42 (Figure 5C and 5D). Thus, the results indicated that PARG has a functional GAP domain which, *in vitro,* is active on Rho, Rac and Cdc42, but with a clear preference for Rho. It is likely, therefore, that Rho is the physiological target of PARG. The name PARG is consequently derived from PTPL1 Associated RhoGAP.

Example 8: PDZ4 Binds to the C-terminal portion of PARG

It has been shown that PDZ domains interact with the C-terminal ends of short peptides and that a valine residue at the absolute C-terminal end is important for binding (Kim *et al.*, 1995; Kornau *et al.*, 1995; Saras *et al.*, in preparation). Since PARG was identified through a specific interaction with PDZ 4 of PTPL1, and since it has a valine residue at the C-terminal end, we found it likely that the interaction is mediated via PDZ 4 and the C-terminal tail of PARG. To verify this possibility, peptides corresponding to the last 4, 5 or 6 C-terminal amino acid residues of PARG (PQFV, IPQFV and EIPQFV; SEQ ID Nos:7, 9 and 11) were synthesized in an Applied Biosystems 430A Peptide Synthesizer using t-butoxycarbonyl chemistry and purified by reversed phase high performance liquid chromatography. The peptides were coupled to Affigel 15 beads (Bio-Rad, Richmond, CA) via their N-terminal ends following the manufacturers instructions and incubated with GST-PDZ fusion proteins (50 nM) at 4°C for 2 h in binding buffer (20 mM Tris-HCl, pH 7.4, 150 mM NaCl, 10 mM EDTA, 0.5% Triton X-100,

0.5% deoxycholate, 1 mM dithiothreitol). The beads were washed four times in binding buffer and bound fusion proteins were eluted by boiling for 5 min in SDS-sample buffer and subjected to SDS-gel electrophoresis using 11 % polyacrylamide gels. After electrophoresis, the proteins were transferred to nitrocellulose membranes (Hybond C Extra; Amersham) and the membranes were incubated with α-GST antiserum (rabbit antiserum raised against recombinant GST expressed in bacteria). Bound antibodies were visualized by using enhanced chemiluminescence (ECL, Amersham), according to the manufacturer's instructions.

As shown in Figure 6, the fusion proteins GST-PDZ 4 and GST-PDZ 4-5, but not GST fusion proteins containing PDZ 1, PDZ 2, PDZ 3 or PDZ 5 only, bound to the peptide corresponding to the last four amino acid residues of PARG. Similar results were obtained by using the longer peptides, indicating that a maximum of four amino acid residues at the C-terminal end of PARG is enough for a strong and specific interaction with PDZ 4 of PTPL1.

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

All references disclosed herein are incorporated by reference in their entirety.

A Sequence Listing is presented below and is followed by what is claimed.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: LUDWIG INSTITUTE FOR CANCER RESEARCH
(B) STREET: 1345 AVENUE OF THE AMERICAS
(C) CITY: NEW YORK
(D) STATE: NEW YORK
(E) COUNTRY: UNITED STATES OF AMERICA
(F) POSTAL CODE: 10105

(ii) TITLE OF INVENTION: PARG, A GTPASE ACTIVATING PROTEIN WHICH INTERACTS WITH PTPL1

(iii) NUMBER OF SEQUENCES: 13

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
(B) STREET: 600 ATLANTIC AVENUE
(C) CITY: BOSTON
(D) STATE: MASSACHUSETTS
(E) COUNTRY: UNITED STATES OF AMERICA
(F) POSTAL CODE: 02210

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/805,583
(B) FILING DATE: 25-FEB-1997

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Van Amsterdam, John R.
(B) REGISTRATION NUMBER: 40,212
(C) REFERENCE/DOCKET NUMBER: L0461/7007WO

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 617-720-3500
(B) TELEFAX: 617-720-2441

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5238 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 184..3966

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
GCTGTGGCTG CGGCTGCGGC TGCGGCTGAG ATTTGGCCGG GCGTCCGCAG GCCGTGGGGG      60

ATGGGGGCAG CGAGCTCCAG CCCTCGGCGG TGGCGGCGGC CGTAGGTGTG GGGCGGGCGT     120

CCGCGTCCGG CACGCGAGAT GGAGCGCCGT GGATTTCAGT TTTTCTGACT GTTACATGAA     180

AGG ATG ATT GCT CAC AAA CAG AAA AAG ACA AAG AAA AAA CGT GCT TGG        228
    Met Ile Ala His Lys Gln Lys Lys Thr Lys Lys Lys Arg Ala Trp
      1               5                  10                  15

GCA TCA GGT CAA CTC TCT ACT GAT ATT ACA ACT TCT GAA ATG GGG CTC        276
Ala Ser Gly Gln Leu Ser Thr Asp Ile Thr Thr Ser Glu Met Gly Leu
                  20                  25                  30

AAG TCC TTA AGT TCC AAC TCT ATT TTT GAT CCG GAT TAC ATC AAG GAG        324
Lys Ser Leu Ser Ser Asn Ser Ile Phe Asp Pro Asp Tyr Ile Lys Glu
              35                  40                  45

TTG GTG AAT GAT ATC AGG AAG TTC TCC CAC ATC TTA CTA TAT TTG AAA        372
Leu Val Asn Asp Ile Arg Lys Phe Ser His Ile Leu Leu Tyr Leu Lys
          50                  55                  60

GAA GCC ATA TTT TCA GAC TGT TTT AAA GAA GTT ATT CAT ATA CGT CTA        420
Glu Ala Ile Phe Ser Asp Cys Phe Lys Glu Val Ile His Ile Arg Leu
      65                  70                  75

GAG GAA CTG CTC CGT GTT TTA AAG TCT ATA ATG AAT AAA CAT CAG AAC        468
Glu Glu Leu Leu Arg Val Leu Lys Ser Ile Met Asn Lys His Gln Asn
  80                  85                  90                  95
```

```
CTC AAT TCT GTT GAT CTT CAA AAT GCT GCA GAA ATG CTC ACT GCA AAA      516
Leu Asn Ser Val Asp Leu Gln Asn Ala Ala Glu Met Leu Thr Ala Lys
                100                 105                 110

GTG AAA GCT GTG AAC TTC ACA GAA GTT AAT GAA GAA AAC AAA AAC GAT      564
Val Lys Ala Val Asn Phe Thr Glu Val Asn Glu Glu Asn Lys Asn Asp
            115                 120                 125

CTC TTC CAG GAA GTG TTT TCT TCT ATT GAA ACT TTG GCA TTT ACC TTT      612
Leu Phe Gln Glu Val Phe Ser Ser Ile Glu Thr Leu Ala Phe Thr Phe
        130                 135                 140

GGA AAT ATC CTT ACA AAC TTC CTT ATG GGA GAT GTA GGC AAT GAT TCA      660
Gly Asn Ile Leu Thr Asn Phe Leu Met Gly Asp Val Gly Asn Asp Ser
    145                 150                 155

TTC TTG CGA CTG CCT GTT TCT CGA GAA ACT AAG TCG TTT GAA AAT GTT      708
Phe Leu Arg Leu Pro Val Ser Arg Glu Thr Lys Ser Phe Glu Asn Val
160                 165                 170                 175

TCT GTG GAA TCA GTG GAC TCA TCC AGT GAA AAA GGA AAT TTT TCC CCT      756
Ser Val Glu Ser Val Asp Ser Ser Ser Glu Lys Gly Asn Phe Ser Pro
                180                 185                 190

TTA GAA CTA GAC AAC GTG CTG TTA AAG AAC ACT GAC TCT ATC GAG CTG      804
Leu Glu Leu Asp Asn Val Leu Leu Lys Asn Thr Asp Ser Ile Glu Leu
            195                 200                 205

GCT TTG TCA TAT GCT AAA ACT TGG TCA AAA TAT ACT AAG AAC ATA GTT      852
Ala Leu Ser Tyr Ala Lys Thr Trp Ser Lys Tyr Thr Lys Asn Ile Val
        210                 215                 220

TCA TGG GTT GAA AAA AAG CTT AAC TTG GAA TTG GAG TCC ACT AGA AAT      900
Ser Trp Val Glu Lys Lys Leu Asn Leu Glu Leu Glu Ser Thr Arg Asn
    225                 230                 235

ATG GTC AAG TTG GCA GAG GCA ACT AGA ACT AAC ATT GGA ATT CAG GAG      948
Met Val Lys Leu Ala Glu Ala Thr Arg Thr Asn Ile Gly Ile Gln Glu
240                 245                 250                 255

TTC ATG CCA CTG CAG TCT CTG TTT ACT AAT GCT CTT CTT AAT GAT ATA      996
Phe Met Pro Leu Gln Ser Leu Phe Thr Asn Ala Leu Leu Asn Asp Ile
                260                 265                 270

GAA AGC AGT CAC CTT TTA CAA CAA ACA ATT GCA GCT CTC CAG GCT AAC     1044
Glu Ser Ser His Leu Leu Gln Gln Thr Ile Ala Ala Leu Gln Ala Asn
            275                 280                 285
```

```
AAA TTT GTG CAG CCT CTA CTT GGA AGG AAA AAT GAA ATG GAA AAA CAA      1092
Lys Phe Val Gln Pro Leu Leu Gly Arg Lys Asn Glu Met Glu Lys Gln
        290                 295                 300

AGG AAA GAA ATA AAA GAG CTT TGG AAA CAG GAG CAA AAT AAA ATG CTT      1140
Arg Lys Glu Ile Lys Glu Leu Trp Lys Gln Glu Gln Asn Lys Met Leu
    305                 310                 315

GAA GCA GAG AAT GCT CTC AAA AAG GCA AAA TTA TTA TGC ATG CAA CGT      1188
Glu Ala Glu Asn Ala Leu Lys Lys Ala Lys Leu Leu Cys Met Gln Arg
320                 325                 330                 335

CAA GAT GAA TAT GAG AAA GCA AAG TCT TCC ATG TTT CGT GCA GAA GAG      1236
Gln Asp Glu Tyr Glu Lys Ala Lys Ser Ser Met Phe Arg Ala Glu Glu
                340                 345                 350

GAG CAT CTG TCT TCA AGT GGC GGA TTA GCA AAA AAT CTC AAC AAG CAA      1284
Glu His Leu Ser Ser Ser Gly Gly Leu Ala Lys Asn Leu Asn Lys Gln
            355                 360                 365

CTA GAA AAA AAG CGA AGG TTG GAA GAG GAG GCT CTC CAA AAA GTA GAA      1332
Leu Glu Lys Lys Arg Arg Leu Glu Glu Glu Ala Leu Gln Lys Val Glu
        370                 375                 380

GAA GCA GAT GAA CTT TAC AAA GTT TGT GTG ACA AAT GTT GAA GAA AGA      1380
Glu Ala Asp Glu Leu Tyr Lys Val Cys Val Thr Asn Val Glu Glu Arg
    385                 390                 395

AGA AAT GAT GTA GAA AAT ACC AAA AGA GAA ATT TTA GCA CAA CTC CGG      1428
Arg Asn Asp Val Glu Asn Thr Lys Arg Glu Ile Leu Ala Gln Leu Arg
400                 405                 410                 415

ACA CTT GTT TTC CAG TGT GAT CTT ACC CTT AAA GCG GTA ACA GTT AAC      1476
Thr Leu Val Phe Gln Cys Asp Leu Thr Leu Lys Ala Val Thr Val Asn
                420                 425                 430

CTC TTC CAC ATG CAG CAT CTG CAG GCT GCT TCC CTT GCA GAC AGA TTA      1524
Leu Phe His Met Gln His Leu Gln Ala Ala Ser Leu Ala Asp Arg Leu
            435                 440                 445

CAG TCT CTC TGT GGT AGT GCC AAA CTC TAT GAC CCA GGC CAA GAG TAC      1572
Gln Ser Leu Cys Gly Ser Ala Lys Leu Tyr Asp Pro Gly Gln Glu Tyr
        450                 455                 460

AGT GAA TTT GTC AAG GCC ACA AAT TCA ACT GAA GAA GAA AAA GTT GAT      1620
Ser Glu Phe Val Lys Ala Thr Asn Ser Thr Glu Glu Glu Lys Val Asp
    465                 470                 475
```

```
GGA AAT GTA AAT AAA CAT TTA AAT AGT TCC CAA CCT TCA GGA TTT GGA      1668
Gly Asn Val Asn Lys His Leu Asn Ser Ser Gln Pro Ser Gly Phe Gly
480                 485                 490                 495

CCT GCC AAC TCT TTA GAG GAT GTT GTA CGC CTT CCT GAC AGT TCT AAT      1716
Pro Ala Asn Ser Leu Glu Asp Val Val Arg Leu Pro Asp Ser Ser Asn
                500                 505                 510

AAA ATT GAA GAG GAC AGA TGC TCT AAC AGT GCA GAT ATA ACA GGT CCT      1764
Lys Ile Glu Glu Asp Arg Cys Ser Asn Ser Ala Asp Ile Thr Gly Pro
            515                 520                 525

TCC TTT ATA AGA TCA TGG ACA TTT GGG ATG TTT AGT GAT TCT GAG AGC      1812
Ser Phe Ile Arg Ser Trp Thr Phe Gly Met Phe Ser Asp Ser Glu Ser
        530                 535                 540

ACT GGA GGG AGC AGC GAA TCT AGA TCT CTG GAT TCA GAA TCT ATA AGT      1860
Thr Gly Gly Ser Ser Glu Ser Arg Ser Leu Asp Ser Glu Ser Ile Ser
    545                 550                 555

CCA GGA GAC TTT CAT CGA AAA CTT CCA CGA ACA CCA TCC AGT GGA ACT      1908
Pro Gly Asp Phe His Arg Lys Leu Pro Arg Thr Pro Ser Ser Gly Thr
560                 565                 570                 575

ATG TCC TCT GCA GAT GAT CTA GAT GAA AGA GAG CCA CCT TCC CCT TCA      1956
Met Ser Ser Ala Asp Asp Leu Asp Glu Arg Glu Pro Pro Ser Pro Ser
                580                 585                 590

GAA ACT GGA CCC AAT TCC CTT GGA ACA TTT AAG AAA ACA TTG ATG TCA      2004
Glu Thr Gly Pro Asn Ser Leu Gly Thr Phe Lys Lys Thr Leu Met Ser
            595                 600                 605

AAG GCA GCT CTC ACA CAC AAG TTT CGC AAA TTG AGA TCC CCC ACG AAA      2052
Lys Ala Ala Leu Thr His Lys Phe Arg Lys Leu Arg Ser Pro Thr Lys
        610                 615                 620

TGT AGG GAT TGT GAA GGC ATT GTA GTG TTC CAA GGT GTT GAA TGT GAA      2100
Cys Arg Asp Cys Glu Gly Ile Val Val Phe Gln Gly Val Glu Cys Glu
    625                 630                 635

GAG TGT CTC CTT GTT TGT CAT CGA AAG TGT TTG GAA AAT TTA GTC ATT      2148
Glu Cys Leu Leu Val Cys His Arg Lys Cys Leu Glu Asn Leu Val Ile
640                 645                 650                 655

ATT TGT GGT CAT CAG AAA CTT CCA GGA AAA ATA CAC TTA TTT GGA GCA      2196
Ile Cys Gly His Gln Lys Leu Pro Gly Lys Ile His Leu Phe Gly Ala
                660                 665                 670
```

```
GAA TTC ACA CTA GTT GCA AAA AAG GAA CCA GAT GGT ATC CCT TTT ATA      2244
Glu Phe Thr Leu Val Ala Lys Lys Glu Pro Asp Gly Ile Pro Phe Ile
            675                 680                 685

CTC AAA ATA TGT GCC TCA GAG ATT GAA AAT AGA GCT TTG TGT CTA CAG      2292
Leu Lys Ile Cys Ala Ser Glu Ile Glu Asn Arg Ala Leu Cys Leu Gln
        690                 695                 700

GGA ATT TAT CGT GTG TGT GGA AAC AAA ATA AAA ACT GAA AAA TTG TGT      2340
Gly Ile Tyr Arg Val Cys Gly Asn Lys Ile Lys Thr Glu Lys Leu Cys
    705                 710                 715

CTA GCT TTG GAA AAT GGT ATG CAC TTG GTA GAT ATT TCA GAA TTT AGT      2388
Leu Ala Leu Glu Asn Gly Met His Leu Val Asp Ile Ser Glu Phe Ser
720                 725                 730                 735

TCA CAT GAT ATC TGT GAC GTC TTG AAA TTA TAC CTT CGG CAG CTC CCA      2436
Ser His Asp Ile Cys Asp Val Leu Lys Leu Tyr Leu Arg Gln Leu Pro
                740                 745                 750

GAA CCA TTT ATT TTA TTT CGA TTG TAC AAG GAA TTT ATA GAC CTT GCA      2484
Glu Pro Phe Ile Leu Phe Arg Leu Tyr Lys Glu Phe Ile Asp Leu Ala
            755                 760                 765

AAA GAG ATC CAA CAT GTA AAT GAA GAA CAA GAG ACA AAA AAG AAT AGT      2532
Lys Glu Ile Gln His Val Asn Glu Glu Gln Glu Thr Lys Lys Asn Ser
        770                 775                 780

CTT GAA GAC AAA AAA TGG CCA AAT ATG TGT ATA GAA ATA AAC CGA ATT      2580
Leu Glu Asp Lys Lys Trp Pro Asn Met Cys Ile Glu Ile Asn Arg Ile
    785                 790                 795

CTT CTA AAA AGC AAA GAC CTT CTA AGA CAA TTG CCA GCA TCA AAT TTT      2628
Leu Leu Lys Ser Lys Asp Leu Leu Arg Gln Leu Pro Ala Ser Asn Phe
800                 805                 810                 815

AAC AGT CTT CAT TTC CTT ATA GTA CAT CTA AAG CGG GTA GTA GAT CAT      2676
Asn Ser Leu His Phe Leu Ile Val His Leu Lys Arg Val Val Asp His
                820                 825                 830

GCA GAA GAA AAC AAG ATG AAC TCC AAA AAC TTG GGG GTG ATA TTT GGA      2724
Ala Glu Glu Asn Lys Met Asn Ser Lys Asn Leu Gly Val Ile Phe Gly
            835                 840                 845

CCA AGT CTC ATT AGG CCA AGG CCA CAA ACT GCT CCT ATC ACC ATC TCC      2772
Pro Ser Leu Ile Arg Pro Arg Pro Gln Thr Ala Pro Ile Thr Ile Ser
        850                 855                 860
```

```
TCC CTT GCA GAG TAT TCA AAT CAA GCA CGC TTG GTA GAG TTT CTC ATT      2820
Ser Leu Ala Glu Tyr Ser Asn Gln Ala Arg Leu Val Glu Phe Leu Ile
    865                 870                 875

ACT TAC TCA CAG AAG ATC TTC GAT GGG TCC CTA CAA CCA CAA GAT GTT      2868
Thr Tyr Ser Gln Lys Ile Phe Asp Gly Ser Leu Gln Pro Gln Asp Val
880                 885                 890                 895

ATG TGT AGC ATA GGT GTT GTT GAT CAA GGC TGT TTT CCA AAG CCT CTG      2916
Met Cys Ser Ile Gly Val Val Asp Gln Gly Cys Phe Pro Lys Pro Leu
                900                 905                 910

TTA TCA CCA GAA GAA AGA GAC ATT GAA CGT TCC ATG AAG TCA CTA TTT      2964
Leu Ser Pro Glu Glu Arg Asp Ile Glu Arg Ser Met Lys Ser Leu Phe
            915                 920                 925

TTT TCT TCA AAG GAA GAT ATC CAT ACT TCA GAG AGT GAA AGC AAA ATT      3012
Phe Ser Ser Lys Glu Asp Ile His Thr Ser Glu Ser Glu Ser Lys Ile
        930                 935                 940

TTT GAA CGA GCT ACA TCA TTT GAG GAA TCA GAA CGC AAG CAA AAT GCG      3060
Phe Glu Arg Ala Thr Ser Phe Glu Glu Ser Glu Arg Lys Gln Asn Ala
    945                 950                 955

TTA GGA AAA TGT GAT GCA TGT CTC AGT GAC AAA GCA CAG TTG CTT CTA      3108
Leu Gly Lys Cys Asp Ala Cys Leu Ser Asp Lys Ala Gln Leu Leu Leu
960                 965                 970                 975

GAC CAA GAG GCT GAA TCA GCA TCC CAA AAG ATA GAA GAT GGT AAA GCC      3156
Asp Gln Glu Ala Glu Ser Ala Ser Gln Lys Ile Glu Asp Gly Lys Ala
                980                 985                 990

CCT AAG CCA CTT TCT CTG AAA TCT GAT AGG TCA ACA AAC AAT GTG GAG      3204
Pro Lys Pro Leu Ser Leu Lys Ser Asp Arg Ser Thr Asn Asn Val Glu
            995                 1000                1005

AGG CAT ACT CCA AGG ACC AAG ATT AGA CCT GTA AGT TTG CCT GTA GAT      3252
Arg His Thr Pro Arg Thr Lys Ile Arg Pro Val Ser Leu Pro Val Asp
        1010                1015                1020

AGA CTA CTT CTT GCA AGT CCT CCT AAT GAG AGA AAT GGC AGA AAT ATG      3300
Arg Leu Leu Leu Ala Ser Pro Pro Asn Glu Arg Asn Gly Arg Asn Met
    1025                1030                1035

GGA AAT GTA AAT TTA GAC AAG TTT TGC AAG AAT CCT GCC TTT GAA GGA      3348
Gly Asn Val Asn Leu Asp Lys Phe Cys Lys Asn Pro Ala Phe Glu Gly
1040                1045                1050                1055
```

GTT AAT AGA AAA GAC GCT GCT ACT ACT GTT TGT TCC AAA TTT AAT GGC 3396
Val Asn Arg Lys Asp Ala Ala Thr Thr Val Cys Ser Lys Phe Asn Gly
1060 1065 1070

TTT GAC CAG CAA ACT CTA CAG AAA ATT CAG GAC AAA CAG TAT GAA CAA 3444
Phe Asp Gln Gln Thr Leu Gln Lys Ile Gln Asp Lys Gln Tyr Glu Gln
1075 1080 1085

AAC AGC CTA ACT GCC AAG ACT ACA ATG ATC ATG CCC AGT GCA CTC CAG 3492
Asn Ser Leu Thr Ala Lys Thr Thr Met Ile Met Pro Ser Ala Leu Gln
1090 1095 1100

GAA AAA GGA GTG ACA ACA AGC CTC CAG ATT AGT GGG GAC CAT TCT ATC 3540
Glu Lys Gly Val Thr Thr Ser Leu Gln Ile Ser Gly Asp His Ser Ile
1105 1110 1115

AAT GCC ACT CAA CCC AGT AAG CCA TAT GCA GAG CCA GTC AGG TCA GTG 3588
Asn Ala Thr Gln Pro Ser Lys Pro Tyr Ala Glu Pro Val Arg Ser Val
1120 1125 1130 1135

AGA GAG GCA TCT GAG AGA CGG TCT TCA GAT TCC TAC CCT CTC GCT CCT 3636
Arg Glu Ala Ser Glu Arg Arg Ser Ser Asp Ser Tyr Pro Leu Ala Pro
1140 1145 1150

GTC AGA GCA CCC AGA ACA CTG CAG CCT CAA CAT TGG ACA ACA TTT TAT 3684
Val Arg Ala Pro Arg Thr Leu Gln Pro Gln His Trp Thr Thr Phe Tyr
1155 1160 1165

AAA CCA CAT GCT CCC ATC ATC AGT ATC AGG GGG AAT GAG GAG AAG CCA 3732
Lys Pro His Ala Pro Ile Ile Ser Ile Arg Gly Asn Glu Glu Lys Pro
1170 1175 1180

GCT TCA CCC TCA GCA GCA TGC CCT CCT GGC ACA GAT CAC GAT CCC CAC 3780
Ala Ser Pro Ser Ala Ala Cys Pro Pro Gly Thr Asp His Asp Pro His
1185 1190 1195

GGT CTC GTG GTG AAG TCA ATG CCA GAC CCA GAC AAA GCA TCA GCT TGT 3828
Gly Leu Val Val Lys Ser Met Pro Asp Pro Asp Lys Ala Ser Ala Cys
1200 1205 1210 1215

CCT GGG CAA GCA ACT GGT CAA CCT AAA GAA GAC TCT GAG GAG CTT GGC 3876
Pro Gly Gln Ala Thr Gly Gln Pro Lys Glu Asp Ser Glu Glu Leu Gly
1220 1225 1230

TTG CCT GAT GTG AAT CCA ATG TGT CAG AGA CCA AGG CTA AAA CGA ATG 3924
Leu Pro Asp Val Asn Pro Met Cys Gln Arg Pro Arg Leu Lys Arg Met
1235 1240 1245

CAA CAG TTT GAA GAC CTC GAA GAT GAA ATT CCA CAA TTT GTG 3966
Gln Gln Phe Glu Asp Leu Glu Asp Glu Ile Pro Gln Phe Val
1250 1255 1260

TAGGGATGTC AAATTTCAGG GTTTTTTTGT TGTTGTTGTG TTATTTTGTG GTATTGTGCT 4026

TGTTTTGTGA AAGAATGTTT TGACAGGGCC CCTTTTGTAT AGGACTGCCA AATCATGGGT 4086

TTTGCCTTTT GTTGTTGTAT TTATCCTCTG TTGGTAATAC TGAATGGTAG AATGTTTTGA 4146

TAGGGTCACA TTTGTGCCTC ACTGGAATTA TCTTTAAATT CTGTATTTTT AAAGTTGTGA 4206

ATAAGATAGG TGGATTCGTA TTTTTTAAAG TTCAGTTGAC TTTCCCCACC AAATGGTCCA 4266

TTTGAATGCA TCCCTAATAT ATGATATAGT CTCAACTAAT AGGTGCAATT TGGGAAAATC 4326

AGGTTTATTT TTTGGAGTGG AACTGTTATA AGTGCTTATT TATAAAAGGA ATGTTTCTGA 4386

ATGCAAGTGC CTAAAAAGAT CTTTGTTGGT ATGCATATGT TTTGTCACAC AATTTATAGT 4446

GCATCTTTCA CCATTTGTGC TTTTTTAAGA TAGTATGTAA GCTCTTATTT TTCAATTGGC 4506

AATTCAGTTA ATTTTTAAAT GTTTACATAA TGGCCAGAAG GCTTGCAAAT CTGTATTTAA 4566

TTGCATTTTA ATTAATTGCC AGTTTTTACA TGTAGTAGTC AGTTGTACAA AGAAAATGCA 4626

CTTAAACCTG TTTCTAAATT ATATATTCAG TTATATTATA TTTGGCTTTA GATGGTTTTA 4686

ATACATTTGA TAGTTTTTCA CCCCTTGGCT TTATTTTATA TAAACTTTTG TTTTTCAGCA 4746

GTTCTGAACT TTTTAGTATT TTATAAATGG TCCAAAAAAT GCCTGTTTCA GAAGTTTTTG 4806

AATTCAGTGC ATTTCCTCTT GATTTGTCTG GGTTAAAACC ATTCCTTTTG TATGAAATGT 4866

TTTGACTTAG GAATCATTTT ATGTACTTGT TCTACCTGGA TTGTCAACAA CTGAAAGTAC 4926

ATATTTCATC CAAATCAAGC TAAAATTTAT TTAAGTTGAT TCTGAGAGTA CAGGTCAGTA 4986

AGCCTCATTA TTTGGAATTT GAGAGAAGTA TAGGTGATCG GATCTGTTTC ATTTATAAAA 5046

GGTCCAGTTT TTAGGACTAG TACATTCCTG TTATTTTCTG GGTTTTATCA TTTTGCCTAA 5106

AATAGGATAT AAAAGGGACA AAAAATAAGT AGACTGTTTT TATGTGTGAA TTATATTTCT 5166

ACTAAATGTT TTTGTATGAC TGTGTTATAC TTGATAATAT ATATATATAT ATATAAAAAA 5226

AAAAAAAAAA AA 5238

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1261 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Ile Ala His Lys Gln Lys Lys Thr Lys Lys Lys Arg Ala Trp Ala
1 5 10 15

Ser Gly Gln Leu Ser Thr Asp Ile Thr Thr Ser Glu Met Gly Leu Lys
20 25 30

Ser Leu Ser Ser Asn Ser Ile Phe Asp Pro Asp Tyr Ile Lys Glu Leu
35 40 45

Val Asn Asp Ile Arg Lys Phe Ser His Ile Leu Leu Tyr Leu Lys Glu
50 55 60

Ala Ile Phe Ser Asp Cys Phe Lys Glu Val Ile His Ile Arg Leu Glu
65 70 75 80

Glu Leu Leu Arg Val Leu Lys Ser Ile Met Asn Lys His Gln Asn Leu
85 90 95

Asn Ser Val Asp Leu Gln Asn Ala Ala Glu Met Leu Thr Ala Lys Val
100 105 110

Lys Ala Val Asn Phe Thr Glu Val Asn Glu Glu Asn Lys Asn Asp Leu
115 120 125

Phe Gln Glu Val Phe Ser Ser Ile Glu Thr Leu Ala Phe Thr Phe Gly
130 135 140

Asn Ile Leu Thr Asn Phe Leu Met Gly Asp Val Gly Asn Asp Ser Phe
145 150 155 160

Leu Arg Leu Pro Val Ser Arg Glu Thr Lys Ser Phe Glu Asn Val Ser
165 170 175

Val Glu Ser Val Asp Ser Ser Ser Glu Lys Gly Asn Phe Ser Pro Leu
180 185 190

Glu Leu Asp Asn Val Leu Leu Lys Asn Thr Asp Ser Ile Glu Leu Ala

195 200 205

Leu Ser Tyr Ala Lys Thr Trp Ser Lys Tyr Thr Lys Asn Ile Val Ser
210 215 220

Trp Val Glu Lys Lys Leu Asn Leu Glu Leu Glu Ser Thr Arg Asn Met
225 230 235 240

Val Lys Leu Ala Glu Ala Thr Arg Thr Asn Ile Gly Ile Gln Glu Phe
245 250 255

Met Pro Leu Gln Ser Leu Phe Thr Asn Ala Leu Leu Asn Asp Ile Glu
260 265 270

Ser Ser His Leu Leu Gln Gln Thr Ile Ala Ala Leu Gln Ala Asn Lys
275 280 285

Phe Val Gln Pro Leu Leu Gly Arg Lys Asn Glu Met Glu Lys Gln Arg
290 295 300

Lys Glu Ile Lys Glu Leu Trp Lys Gln Glu Gln Asn Lys Met Leu Glu
305 310 315 320

Ala Glu Asn Ala Leu Lys Lys Ala Lys Leu Leu Cys Met Gln Arg Gln
325 330 335

Asp Glu Tyr Glu Lys Ala Lys Ser Ser Met Phe Arg Ala Glu Glu Glu
340 345 350

His Leu Ser Ser Ser Gly Gly Leu Ala Lys Asn Leu Asn Lys Gln Leu
355 360 365

Glu Lys Lys Arg Arg Leu Glu Glu Glu Ala Leu Gln Lys Val Glu Glu
370 375 380

Ala Asp Glu Leu Tyr Lys Val Cys Val Thr Asn Val Glu Glu Arg Arg
385 390 395 400

Asn Asp Val Glu Asn Thr Lys Arg Glu Ile Leu Ala Gln Leu Arg Thr
405 410 415

Leu Val Phe Gln Cys Asp Leu Thr Leu Lys Ala Val Thr Val Asn Leu
420 425 430

Phe His Met Gln His Leu Gln Ala Ala Ser Leu Ala Asp Arg Leu Gln
435 440 445

Ser Leu Cys Gly Ser Ala Lys Leu Tyr Asp Pro Gly Gln Glu Tyr Ser

450 455 460

Glu Phe Val Lys Ala Thr Asn Ser Thr Glu Glu Glu Lys Val Asp Gly
465 470 475 480

Asn Val Asn Lys His Leu Asn Ser Ser Gln Pro Ser Gly Phe Gly Pro
485 490 495

Ala Asn Ser Leu Glu Asp Val Val Arg Leu Pro Asp Ser Ser Asn Lys
500 505 510

Ile Glu Glu Asp Arg Cys Ser Asn Ser Ala Asp Ile Thr Gly Pro Ser
515 520 525

Phe Ile Arg Ser Trp Thr Phe Gly Met Phe Ser Asp Ser Glu Ser Thr
530 535 540

Gly Gly Ser Ser Glu Ser Arg Ser Leu Asp Ser Glu Ser Ile Ser Pro
545 550 555 560

Gly Asp Phe His Arg Lys Leu Pro Arg Thr Pro Ser Ser Gly Thr Met
565 570 575

Ser Ser Ala Asp Asp Leu Asp Glu Arg Glu Pro Pro Ser Pro Ser Glu
580 585 590

Thr Gly Pro Asn Ser Leu Gly Thr Phe Lys Lys Thr Leu Met Ser Lys
595 600 605

Ala Ala Leu Thr His Lys Phe Arg Lys Leu Arg Ser Pro Thr Lys Cys
610 615 620

Arg Asp Cys Glu Gly Ile Val Val Phe Gln Gly Val Glu Cys Glu Glu
625 630 635 640

Cys Leu Leu Val Cys His Arg Lys Cys Leu Glu Asn Leu Val Ile Ile
645 650 655

Cys Gly His Gln Lys Leu Pro Gly Lys Ile His Leu Phe Gly Ala Glu
660 665 670

Phe Thr Leu Val Ala Lys Lys Glu Pro Asp Gly Ile Pro Phe Ile Leu
675 680 685

Lys Ile Cys Ala Ser Glu Ile Glu Asn Arg Ala Leu Cys Leu Gln Gly
690 695 700

Ile Tyr Arg Val Cys Gly Asn Lys Ile Lys Thr Glu Lys Leu Cys Leu

705 710 715 720

Ala Leu Glu Asn Gly Met His Leu Val Asp Ile Ser Glu Phe Ser Ser
725 730 735

His Asp Ile Cys Asp Val Leu Lys Leu Tyr Leu Arg Gln Leu Pro Glu
740 745 750

Pro Phe Ile Leu Phe Arg Leu Tyr Lys Glu Phe Ile Asp Leu Ala Lys
755 760 765

Glu Ile Gln His Val Asn Glu Glu Gln Glu Thr Lys Lys Asn Ser Leu
770 775 780

Glu Asp Lys Lys Trp Pro Asn Met Cys Ile Glu Ile Asn Arg Ile Leu
785 790 795 800

Leu Lys Ser Lys Asp Leu Leu Arg Gln Leu Pro Ala Ser Asn Phe Asn
805 810 815

Ser Leu His Phe Leu Ile Val His Leu Lys Arg Val Val Asp His Ala
820 825 830

Glu Glu Asn Lys Met Asn Ser Lys Asn Leu Gly Val Ile Phe Gly Pro
835 840 845

Ser Leu Ile Arg Pro Arg Pro Gln Thr Ala Pro Ile Thr Ile Ser Ser
850 855 860

Leu Ala Glu Tyr Ser Asn Gln Ala Arg Leu Val Glu Phe Leu Ile Thr
865 870 875 880

Tyr Ser Gln Lys Ile Phe Asp Gly Ser Leu Gln Pro Gln Asp Val Met
885 890 895

Cys Ser Ile Gly Val Val Asp Gln Gly Cys Phe Pro Lys Pro Leu Leu
900 905 910

Ser Pro Glu Glu Arg Asp Ile Glu Arg Ser Met Lys Ser Leu Phe Phe
915 920 925

Ser Ser Lys Glu Asp Ile His Thr Ser Glu Ser Glu Ser Lys Ile Phe
930 935 940

Glu Arg Ala Thr Ser Phe Glu Glu Ser Glu Arg Lys Gln Asn Ala Leu
945 950 955 960

Gly Lys Cys Asp Ala Cys Leu Ser Asp Lys Ala Gln Leu Leu Leu Asp

965 970 975

Gln Glu Ala Glu Ser Ala Ser Gln Lys Ile Glu Asp Gly Lys Ala Pro
980 985 990

Lys Pro Leu Ser Leu Lys Ser Asp Arg Ser Thr Asn Asn Val Glu Arg
995 1000 1005

His Thr Pro Arg Thr Lys Ile Arg Pro Val Ser Leu Pro Val Asp Arg
1010 1015 1020

Leu Leu Leu Ala Ser Pro Pro Asn Glu Arg Asn Gly Arg Asn Met Gly
1025 1030 1035 1040

Asn Val Asn Leu Asp Lys Phe Cys Lys Asn Pro Ala Phe Glu Gly Val
1045 1050 1055

Asn Arg Lys Asp Ala Ala Thr Thr Val Cys Ser Lys Phe Asn Gly Phe
1060 1065 1070

Asp Gln Gln Thr Leu Gln Lys Ile Gln Asp Lys Gln Tyr Glu Gln Asn
1075 1080 1085

Ser Leu Thr Ala Lys Thr Thr Met Ile Met Pro Ser Ala Leu Gln Glu
1090 1095 1100

Lys Gly Val Thr Thr Ser Leu Gln Ile Ser Gly Asp His Ser Ile Asn
1105 1110 1115 1120

Ala Thr Gln Pro Ser Lys Pro Tyr Ala Glu Pro Val Arg Ser Val Arg
1125 1130 1135

Glu Ala Ser Glu Arg Arg Ser Ser Asp Ser Tyr Pro Leu Ala Pro Val
1140 1145 1150

Arg Ala Pro Arg Thr Leu Gln Pro Gln His Trp Thr Thr Phe Tyr Lys
1155 1160 1165

Pro His Ala Pro Ile Ile Ser Ile Arg Gly Asn Glu Glu Lys Pro Ala
1170 1175 1180

Ser Pro Ser Ala Ala Cys Pro Pro Gly Thr Asp His Asp Pro His Gly
1185 1190 1195 1200

Leu Val Val Lys Ser Met Pro Asp Pro Asp Lys Ala Ser Ala Cys Pro
1205 1210 1215

Gly Gln Ala Thr Gly Gln Pro Lys Glu Asp Ser Glu Glu Leu Gly Leu

1220 1225 1230

Pro Asp Val Asn Pro Met Cys Gln Arg Pro Arg Leu Lys Arg Met Gln
1235 1240 1245

Gln Phe Glu Asp Leu Glu Asp Glu Ile Pro Gln Phe Val
1250 1255 1260

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 251 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

TTAATAGAAA AGACGCTGCT ACTACTGTTT GTTCCAAATT TAATGGCTTT GACCAGCAAA 60

CTCTACAGAA AATTCAGGAC AAACAGTATG AACAAAACAG CCTAACTGCC AAGACTACAA 120

TGATCATGCC CAGTGCACTC CAGGAAAAAG GAGTGACAAC AAGCCTCCAG ATTAGTGGGG 180

ACCATTCTAT CAATGCCACT NAACCCAGTA AGCCATATGC AGAGCCAGTC AGGTCAGTGA 240

GAGAGGCATC T 251

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 256 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CGGTAAGCCA AGCTCCTCAG AGTCTTCTTT AGGTTNACCA GTTGCTTGCC CAGGACAAGC 60

TGATGCTTTG TCTGGGTCTG GCATTGACTT CACCACGAGA CCGTGGGGAT CGTGATCTGT 120

GCCAGGAGGC ACTGCTGCTG AGGGTGAAGC TGGCTTCTCC TCATTCCCCC TGATACTGAT 180

GATGGGAGCA TGTGGTTTAT AAAATGTTGT CCAATGTTGA GGCTGCAGTG TTCTGGGTGC 240

TCTGACAGGA GCGAGA 256

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 298 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CTTTCTGTGA TAGTGCCAAA CTCTATGACC CAGGCCAAGA GTACAGTGAA TTTGTCAAGG 60

CCACAAATTC AACTGAAGAA GAAAAAGTTG ATGGAAATGT AAATAAACAT TTAAATAGTT 120

CCCAACCTTC AGGATTTGGN CCTGCCAACT CTTTAGAGGA TGTTGTACGC CTTCCTGACA 180

GTTCTAATAA AATTGAAGAG GACAGATGCT CTAACAGTGC AGNTATAACA GGTCCTTCCT 240

TTATAAGATC ATGGACATTT GGGATGTTTA GTGATTCTGA GAGCACTGGA GGGAGCAG 298

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CCACAATTTG TG 12

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Pro Gln Phe Val
1

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

ATTCCACAAT TTGTG 15

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Ile Pro Gln Phe Val

1 5

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GAAATTCCAC AATTTGTG 18

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Glu Ile Pro Gln Phe Val
1 5

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2466 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met His Val Ser Leu Ala Glu Ala Leu Glu Val Arg Gly Gly Pro Leu
1 5 10 15

Gln Glu Glu Glu Ile Trp Ala Val Leu Asn Gln Ser Ala Glu Ser Leu
20 25 30

Gln Glu Leu Phe Arg Lys Val Ser Leu Ala Asp Pro Ala Ala Leu Gly
35 40 45

Phe Ile Ile Ser Pro Trp Ser Leu Leu Leu Leu Pro Ser Gly Ser Val
50 55 60

Ser Phe Thr Asp Glu Asn Ile Ser Asn Gln Asp Leu Arg Ala Phe Thr
65 70 75 80

Ala Pro Glu Val Leu Gln Asn Gln Ser Leu Thr Ser Leu Ser Asp Val
85 90 95

Glu Lys Ile His Ile Tyr Ser Leu Gly Met Thr Leu Tyr Trp Gly Ala
100 105 110

Asp Tyr Glu Val Pro Gln Ser Gln Pro Ile Lys Leu Gly Asp His Leu
115 120 125

Asn Ser Ile Leu Leu Gly Met Cys Glu Asp Val Ile Tyr Ala Arg Val
130 135 140

Ser Val Arg Thr Val Leu Asp Ala Cys Ser Ala His Ile Arg Asn Ser
145 150 155 160

Asn Cys Ala Pro Ser Phe Ser Tyr Val Lys His Leu Val Lys Leu Val
165 170 175

Leu Gly Asn Leu Ser Gly Thr Asp Gln Leu Ser Cys Asn Ser Glu Gln
180 185 190

Lys Pro Asp Arg Ser Gln Ala Ile Arg Asp Arg Leu Arg Gly Lys Gly
195 200 205

Leu Pro Thr Gly Arg Ser Ser Thr Ser Asp Val Leu Asp Ile Gln Lys
210 215 220

Pro Pro Leu Ser His Gln Thr Phe Leu Asn Lys Gly Leu Ser Lys Ser
225 230 235 240

Met Gly Phe Leu Ser Ile Lys Asp Thr Gln Asp Glu Asn Tyr Phe Lys

245 250 255

Asp Ile Leu Ser Asp Asn Ser Gly Arg Glu Asp Ser Glu Asn Thr Phe
260 265 270

Ser Pro Tyr Gln Phe Lys Thr Ser Gly Pro Glu Lys Lys Pro Ile Pro
275 280 285

Gly Ile Asp Val Leu Ser Lys Lys Lys Ile Trp Ala Ser Ser Met Asp
290 295 300

Leu Leu Cys Thr Ala Asp Arg Asp Phe Ser Ser Gly Glu Thr Ala Thr
305 310 315 320

Tyr Arg Arg Cys His Pro Glu Ala Val Thr Val Arg Thr Ser Thr Thr
325 330 335

Pro Arg Lys Lys Glu Ala Arg Tyr Ser Asp Gly Ser Ile Ala Leu Asp
340 345 350

Ile Phe Gly Pro Gln Lys Met Asp Pro Ile Tyr His Thr Arg Glu Leu
355 360 365

Pro Thr Ser Ser Ala Ile Ser Ser Ala Leu Asp Arg Ile Arg Glu Arg
370 375 380

Gln Lys Lys Leu Gln Val Leu Arg Glu Ala Met Asn Val Glu Glu Pro
385 390 395 400

Val Arg Arg Tyr Lys Thr Tyr His Gly Asp Val Phe Ser Thr Ser Ser
405 410 415

Glu Ser Pro Ser Ile Ile Ser Ser Glu Ser Asp Phe Arg Gln Val Arg
420 425 430

Arg Ser Glu Ala Ser Lys Arg Phe Glu Ser Ser Ser Gly Leu Pro Gly
435 440 445

Val Asp Glu Thr Leu Ser Gln Gly Gln Ser Gln Arg Pro Ser Arg Gln
450 455 460

Tyr Glu Thr Pro Phe Glu Gly Asn Leu Ile Asn Gln Glu Ile Met Leu
465 470 475 480

Lys Arg Gln Glu Glu Glu Leu Met Gln Leu Gln Ala Lys Met Ala Leu
485 490 495

Arg Gln Ser Arg Leu Ser Leu Tyr Pro Gly Asp Thr Ile Lys Ala Ser

500 505 510

Met Leu Asp Ile Thr Arg Asp Pro Leu Arg Glu Ile Ala Leu Glu Thr
515 520 525

Ala Met Thr Gln Arg Lys Leu Arg Asn Phe Phe Gly Pro Glu Phe Val
530 535 540

Lys Met Thr Ile Glu Pro Phe Ile Ser Leu Asp Leu Pro Arg Ser Ile
545 550 555 560

Leu Thr Lys Lys Gly Lys Asn Glu Asp Asn Arg Arg Lys Val Asn Ile
565 570 575

Met Leu Leu Asn Gly Gln Arg Leu Glu Leu Thr Cys Asp Thr Lys Thr
580 585 590

Ile Cys Lys Asp Val Phe Asp Met Val Val Ala His Ile Gly Leu Val
595 600 605

Glu His His Leu Phe Ala Leu Ala Thr Leu Lys Asp Asn Glu Tyr Phe
610 615 620

Phe Val Asp Pro Asp Leu Lys Leu Thr Lys Val Ala Pro Glu Gly Trp
625 630 635 640

Lys Glu Glu Pro Lys Lys Lys Thr Lys Ala Thr Val Asn Phe Thr Leu
645 650 655

Phe Phe Arg Ile Lys Phe Phe Met Asp Asp Val Ser Leu Ile Gln His
660 665 670

Thr Leu Thr Cys His Gln Tyr Tyr Leu Gln Leu Arg Lys Asp Ile Leu
675 680 685

Glu Glu Arg Met His Cys Asp Asp Glu Thr Ser Leu Leu Leu Ala Ser
690 695 700

Leu Ala Leu Gln Ala Glu Tyr Gly Asp Tyr Gln Pro Glu Val His Gly
705 710 715 720

Val Ser Tyr Phe Arg Met Glu His Tyr Leu Pro Ala Arg Val Met Glu
725 730 735

Lys Leu Asp Leu Ser Tyr Ile Lys Glu Glu Leu Pro Lys Leu His Asn
740 745 750

Thr Tyr Val Gly Ala Ser Glu Lys Glu Thr Glu Leu Glu Phe Leu Lys

755 760 765

Val Cys Gln Arg Leu Thr Glu Tyr Gly Val His Phe His Arg Val His
770 775 780

Pro Glu Lys Lys Ser Gln Thr Gly Ile Leu Leu Gly Val Cys Ser Lys
785 790 795 800

Gly Val Leu Val Phe Glu Val His Asn Gly Val Arg Thr Leu Val Leu
805 810 815

Arg Phe Pro Trp Arg Glu Thr Lys Lys Ile Ser Phe Ser Lys Lys Lys
820 825 830

Ile Thr Leu Gln Asn Thr Ser Asp Gly Ile Lys His Gly Phe Gln Thr
835 840 845

Asp Asn Ser Lys Ile Cys Gln Tyr Leu Leu His Leu Cys Ser Tyr Gln
850 855 860

His Lys Phe Gln Leu Gln Met Arg Ala Arg Gln Ser Asn Gln Asp Ala
865 870 875 880

Gln Asp Ile Glu Arg Ala Ser Phe Arg Ser Leu Asn Leu Gln Ala Glu
885 890 895

Ser Val Arg Gly Phe Asn Met Gly Arg Ala Ile Ser Thr Gly Ser Leu
900 905 910

Ala Ser Ser Thr Leu Asn Lys Leu Ala Val Arg Pro Leu Ser Val Gln
915 920 925

Ala Glu Ile Leu Lys Arg Leu Ser Cys Ser Glu Leu Ser Leu Tyr Gln
930 935 940

Pro Leu Gln Asn Ser Ser Lys Glu Lys Asn Asp Lys Ala Ser Trp Glu
945 950 955 960

Glu Lys Pro Arg Glu Met Ser Lys Ser Tyr His Asp Leu Ser Gln Ala
965 970 975

Ser Leu Tyr Pro His Arg Lys Asn Val Ile Val Asn Met Glu Pro Pro
980 985 990

Pro Gln Thr Val Ala Glu Leu Val Gly Lys Pro Ser His Gln Met Ser
995 1000 1005

Arg Ser Asp Ala Glu Ser Leu Ala Gly Val Thr Lys Leu Asn Asn Ser

1010 1015 1020

Lys Ser Val Ala Ser Leu Asn Arg Ser Pro Glu Arg Arg Lys His Glu
1025 1030 1035 1040

Ser Asp Ser Ser Ser Ile Glu Asp Pro Gly Gln Ala Tyr Val Leu Asp
1045 1050 1055

Val Leu His Lys Arg Trp Ser Ile Val Ser Ser Pro Glu Arg Glu Ile
1060 1065 1070

Thr Leu Val Asn Leu Lys Lys Asp Ala Lys Tyr Gly Leu Gly Phe Gln
1075 1080 1085

Ile Ile Gly Gly Glu Lys Met Gly Arg Leu Asp Leu Gly Ile Phe Ile
1090 1095 1100

Ser Ser Val Ala Pro Gly Gly Pro Ala Asp Phe His Gly Cys Leu Lys
1105 1110 1115 1120

Pro Gly Asp Arg Leu Ile Ser Val Asn Ser Val Ser Leu Glu Gly Val
1125 1130 1135

Ser His His Ala Ala Ile Glu Ile Leu Gln Asn Ala Pro Glu Asp Val
1140 1145 1150

Thr Leu Val Ile Ser Gln Pro Lys Glu Lys Ile Ser Lys Val Pro Ser
1155 1160 1165

Thr Pro Val His Leu Thr Asn Glu Met Lys Asn Tyr Met Lys Lys Ser
1170 1175 1180

Ser Tyr Met Gln Asp Ser Ala Ile Asp Ser Ser Ser Lys Asp His His
1185 1190 1195 1200

Trp Ser Arg Gly Thr Leu Arg His Ile Ser Glu Asn Ser Phe Gly Pro
1205 1210 1215

Ser Gly Gly Leu Arg Glu Gly Ser Leu Ser Ser Gln Asp Ser Arg Thr
1220 1225 1230

Glu Ser Ala Ser Leu Ser Gln Ser Gln Val Asn Gly Phe Phe Ala Ser
1235 1240 1245

His Leu Gly Asp Gln Thr Trp Gln Glu Ser Gln His Gly Ser Pro Ser
1250 1255 1260

Pro Ser Val Ile Ser Lys Ala Thr Glu Lys Glu Thr Phe Thr Asp Ser

1265 1270 1275 1280

Asn Gln Ser Lys Thr Lys Lys Pro Gly Ile Ser Asp Val Thr Asp Tyr
1285 1290 1295

Ser Asp Arg Gly Asp Ser Asp Met Asp Glu Ala Thr Tyr Ser Ser Ser
1300 1305 1310

Gln Asp His Gln Thr Pro Lys Gln Glu Ser Ser Ser Ser Val Asn Thr
1315 1320 1325

Ser Asn Lys Met Asn Phe Lys Thr Phe Ser Ser Ser Pro Pro Lys Pro
1330 1335 1340

Gly Asp Ile Phe Glu Val Glu Leu Ala Lys Asn Asp Asn Ser Leu Gly
1345 1350 1355 1360

Ile Ser Val Thr Gly Gly Val Asn Thr Ser Val Arg His Gly Gly Ile
1365 1370 1375

Tyr Val Lys Ala Val Ile Pro Gln Gly Ala Ala Glu Ser Asp Gly Arg
1380 1385 1390

Ile His Lys Gly Asp Arg Val Leu Ala Val Asn Gly Val Ser Leu Glu
1395 1400 1405

Gly Ala Thr His Lys Gln Ala Val Glu Thr Leu Arg Asn Thr Gly Gln
1410 1415 1420

Val Val His Leu Leu Leu Glu Lys Gly Gln Ser Pro Thr Ser Lys Glu
1425 1430 1435 1440

His Val Pro Val Thr Pro Gln Cys Thr Leu Ser Asp Gln Asn Ala Gln
1445 1450 1455

Gly Gln Gly Pro Glu Lys Val Lys Lys Thr Thr Gln Val Lys Asp Tyr
1460 1465 1470

Ser Phe Val Thr Glu Glu Asn Thr Phe Glu Val Lys Leu Phe Lys Asn
1475 1480 1485

Ser Ser Gly Leu Gly Phe Ser Phe Ser Arg Glu Asp Asn Leu Ile Pro
1490 1495 1500

Glu Gln Ile Asn Ala Ser Ile Val Arg Val Lys Lys Leu Phe Ala Gly
1505 1510 1515 1520

Gln Pro Ala Ala Glu Ser Gly Lys Ile Asp Val Gly Asp Val Ile Leu

1525 1530 1535

Lys Val Asn Gly Ala Ser Leu Lys Gly Leu Ser Gln Gln Glu Val Ile
1540 1545 1550

Ser Ala Leu Arg Gly Thr Ala Pro Glu Val Phe Leu Leu Leu Cys Arg
1555 1560 1565

Pro Pro Pro Gly Val Leu Pro Glu Ile Asp Thr Ala Leu Leu Thr Pro
1570 1575 1580

Leu Gln Ser Pro Ala Gln Val Leu Pro Asn Ser Ser Lys Asp Ser Ser
1585 1590 1595 1600

Gln Pro Ser Cys Val Glu Gln Ser Thr Ser Ser Asp Glu Asn Glu Met
1605 1610 1615

Ser Asp Lys Ser Lys Lys Gln Cys Lys Ser Pro Ser Arg Arg Asp Ser
1620 1625 1630

Tyr Ser Asp Ser Ser Gly Ser Gly Glu Asp Asp Leu Val Thr Ala Pro
1635 1640 1645

Ala Asn Ile Ser Asn Ser Thr Trp Ser Ser Ala Leu His Gln Thr Leu
1650 1655 1660

Ser Asn Met Val Ser Gln Ala Gln Ser His His Glu Ala Pro Lys Ser
1665 1670 1675 1680

Gln Glu Asp Thr Ile Cys Thr Met Phe Tyr Tyr Pro Gln Lys Ile Pro
1685 1690 1695

Asn Lys Pro Glu Phe Glu Asp Ser Asn Pro Ser Pro Leu Pro Pro Asp
1700 1705 1710

Met Ala Pro Gly Gln Ser Tyr Gln Pro Gln Ser Glu Ser Ala Ser Ser
1715 1720 1725

Ser Ser Met Asp Lys Tyr His Ile His His Ile Ser Glu Pro Thr Arg
1730 1735 1740

Gln Glu Asn Trp Thr Pro Leu Lys Asn Asp Leu Glu Asn His Leu Glu
1745 1750 1755 1760

Asp Phe Glu Leu Glu Val Glu Leu Leu Ile Thr Leu Ile Lys Ser Glu
1765 1770 1775

Lys Ala Ser Leu Gly Phe Thr Val Thr Lys Gly Asn Gln Arg Ile Gly

1780 1785 1790

Cys Tyr Val His Asp Val Ile Gln Asp Pro Ala Lys Ser Asp Gly Arg
1795 1800 1805

Leu Lys Pro Gly Asp Arg Leu Ile Lys Val Asn Asp Thr Asp Val Thr
1810 1815 1820

Asn Met Thr His Thr Asp Ala Val Asn Leu Leu Arg Ala Ala Ser Lys
1825 1830 1835 1840

Thr Val Arg Leu Val Ile Gly Arg Val Leu Glu Leu Pro Arg Ile Pro
1845 1850 1855

Met Leu Pro His Leu Leu Pro Asp Ile Thr Leu Thr Cys Asn Lys Glu
1860 1865 1870

Glu Leu Gly Phe Ser Leu Cys Gly Gly His Asp Ser Leu Tyr Gln Val
1875 1880 1885

Val Tyr Ile Ser Asp Ile Asn Pro Arg Ser Val Ala Ala Ile Glu Gly
1890 1895 1900

Asn Leu Gln Leu Leu Asp Val Ile His Tyr Val Asn Gly Val Ser Thr
1905 1910 1915 1920

Gln Gly Met Thr Leu Glu Glu Val Asn Arg Ala Leu Asp Met Ser Leu
1925 1930 1935

Pro Ser Leu Val Leu Lys Ala Thr Arg Asn Asp Leu Pro Val Val Pro
1940 1945 1950

Ser Ser Lys Arg Ser Ala Val Ser Ala Pro Lys Ser Thr Lys Gly Asn
1955 1960 1965

Gly Ser Tyr Ser Val Gly Ser Cys Ser Gln Pro Ala Leu Thr Pro Asn
1970 1975 1980

Asp Ser Phe Ser Thr Val Ala Gly Glu Glu Ile Asn Glu Ile Ser Tyr
1985 1990 1995 2000

Pro Lys Gly Lys Cys Ser Thr Tyr Gln Ile Lys Gly Ser Pro Asn Leu
2005 2010 2015

Thr Leu Pro Lys Glu Ser Tyr Ile Gln Glu Asp Asp Ile Tyr Asp Asp
2020 2025 2030

Ser Gln Glu Ala Glu Val Ile Gln Ser Leu Leu Asp Val Val Asp Glu

2035 2040 2045

Glu Ala Gln Asn Leu Leu Asn Glu Asn Asn Ala Ala Gly Tyr Ser Cys
2050 2055 2060

Gly Pro Gly Thr Leu Lys Met Asn Gly Lys Leu Ser Glu Glu Arg Thr
2065 2070 2075 2080

Glu Asp Thr Asp Cys Asp Gly Ser Pro Leu Pro Glu Tyr Phe Thr Glu
2085 2090 2095

Ala Thr Lys Met Asn Gly Cys Glu Glu Tyr Cys Glu Glu Lys Val Lys
2100 2105 2110

Ser Glu Ser Leu Ile Gln Lys Pro Gln Glu Lys Lys Thr Asp Asp Asp
2115 2120 2125

Glu Ile Thr Trp Gly Asn Asp Glu Leu Pro Ile Glu Arg Thr Asn His
2130 2135 2140

Glu Asp Ser Asp Lys Asp His Ser Phe Leu Thr Asn Asp Glu Leu Ala
2145 2150 2155 2160

Val Leu Pro Val Val Lys Val Leu Pro Ser Gly Lys Tyr Thr Gly Ala
2165 2170 2175

Asn Leu Lys Ser Val Ile Arg Val Leu Arg Gly Leu Leu Asp Gln Gly
2180 2185 2190

Ile Pro Ser Lys Glu Leu Glu Asn Leu Gln Glu Leu Lys Pro Leu Asp
2195 2200 2205

Gln Cys Leu Ile Gly Gln Thr Lys Glu Asn Arg Arg Lys Asn Arg Tyr
2210 2215 2220

Lys Asn Ile Leu Pro Tyr Asp Ala Thr Arg Val Pro Leu Gly Asp Glu
2225 2230 2235 2240

Gly Gly Tyr Ile Asn Ala Ser Phe Ile Lys Ile Pro Val Gly Lys Glu
2245 2250 2255

Glu Phe Val Tyr Ile Ala Cys Gln Gly Pro Leu Pro Thr Thr Val Gly
2260 2265 2270

Asp Phe Trp Gln Met Ile Trp Glu Gln Lys Ser Thr Val Ile Ala Met
2275 2280 2285

Met Thr Gln Glu Val Glu Gly Glu Lys Ile Lys Cys Gln Arg Tyr Trp

2290 2295 2300

Pro Asn Ile Leu Gly Lys Thr Thr Met Val Ser Asn Arg Leu Arg Leu
2305 2310 2315 2320

Ala Leu Val Arg Met Gln Gln Leu Lys Gly Phe Val Val Arg Ala Met
2325 2330 2335

Thr Leu Glu Asp Ile Gln Thr Arg Glu Val Arg His Ile Ser His Leu
2340 2345 2350

Asn Phe Thr Ala Trp Pro Asp His Asp Thr Pro Ser Gln Pro Asp Asp
2355 2360 2365

Leu Leu Thr Phe Ile Ser Tyr Met Arg His Ile His Arg Ser Gly Pro
2370 2375 2380

Ile Ile Thr His Cys Ser Ala Gly Ile Gly Arg Ser Gly Thr Leu Ile
2385 2390 2395 2400

Cys Ile Asp Val Val Leu Gly Leu Ile Ser Gln Asp Leu Asp Phe Asp
2405 2410 2415

Ile Ser Asp Leu Val Arg Cys Met Arg Leu Gln Arg His Gly Met Val
2420 2425 2430

Gln Thr Glu Asp Gln Tyr Ile Phe Cys Tyr Gln Val Ile Leu Tyr Val
2435 2440 2445

Leu Thr Arg Leu Gln Ala Glu Glu Glu Gln Lys Gln Gln Pro Gln Leu
2450 2455 2460

Leu Lys
2465

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 261 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CTACTGTTTG TTCCAAATTT AATGGCTTTG ACCAGCAAAC TCTACAGAAA ATTCAGGACA 60

AACAGTATGA ACAAAACAGC CTAACTGCCA AGACTACAAT GATCATGCCC AGTGCACTCC 120

AGGAAAAAGG AGTGACAACA AGCCTCCAGA TTAGTGGGGA CCATTCTATC AATGCCACTC 180

AACCCAGTAA GCCATATGCA GAGCCAGTCA GGTCAGTNAG AGAGGCATCT GAGAGACGGT 240

CTTCAGATTC CTACCCTCTC G 261

CLAIMS

1. An isolated nucleic acid molecule selected from the group consisting of

(a) nucleic acid molecules which hybridize under stringent conditions to a molecule consisting of the nucleic acid sequence of SEQ ID NO:1 and which codes for a GTPase-activating polypeptide,

(b) nucleic acid molecules that differ from the nucleic acid molecules of (a) in codon sequence due to the degeneracy of the genetic code, and

(c) complements of (a) and (b).

2. The isolated nucleic acid molecule of claim 1, wherein the isolated nucleic acid molecule comprises nucleotides 184-3966 of SEQ ID NO:1.

3. The isolated nucleic acid molecule of claim 1, wherein the isolated nucleic acid molecule consists essentially of SEQ ID NO:1.

4. The isolated nucleic acid molecule of claim 1, wherein the isolated nucleic acid molecule comprises a molecule having a sequence which encodes amino acids 658-898 of SEQ ID NO:2.

5. An isolated nucleic acid molecule selected from the group consisting of (a) a unique fragment of nucleotides 184-3966 of SEQ ID NO:1 between 12 and 3781 nucleotides in length and (b) complements of "(a)", provided that the nucleic acid molecule excludes sequences consisting only of SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5 and SEQ ID NO:13.

6. The isolated nucleic acid molecule of claim 5, wherein the isolated nucleic acid molecule consists of at least 14 contiguous nucleotides.

7. The isolated nucleic acid molecule of claim 5, wherein the isolated nucleic acid molecule consists of at least 15 contiguous nucleotides.

8. The isolated nucleic acid molecule of claim 5, wherein the isolated nucleic acid molecule consists of at least 16 contiguous nucleotides.

9. The isolated nucleic acid molecule of claim 5, wherein the isolated nucleic acid molecule consists of at least 17 contiguous nucleotides.

10. The isolated nucleic acid molecule of claim 5, wherein the isolated nucleic acid molecule consists of at least 18 contiguous nucleotides.

11. The isolated nucleic acid molecule of claim 5, wherein the isolated nucleic acid molecule consists of at least 20 contiguous nucleotides.

12. The isolated nucleic acid molecule of claim 5, wherein the isolated nucleic acid molecule consists of at least 22 contiguous nucleotides.

13. The isolated nucleic acid molecule of claim 5, wherein the isolated nucleic acid molecule consists of between 12 and 32 contiguous nucleotides.

14. An isolated nucleic acid molecule selected from the group consisting of

(a) a molecule having a sequence selected from the group consisting of SEQ ID NO:6, SEQ ID NO:8 and SEQ ID NO:10;

(b) nucleic acid molecules that differ from the nucleic acid molecules of (a) in codon sequence due to the degeneracy of the genetic code; and

(c) complements of (a) and (b).

15. An expression vector comprising the isolated nucleic acid molecule of claims 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, or 14 operably linked to a promoter.

16. A host cell transformed or transfected with the expression vector of claim 15.

17. An isolated polypeptide encoded by the isolated nucleic acid molecule of claim 1, 2 or 4, wherein the polypeptide has GTPase activating activity.

18. The isolated polypeptide of claim 17, wherein the isolated polypeptide comprises a polypeptide having the sequence of amino acids 658-898 of SEQ ID NO:2.

19. An isolated polypeptide comprising a polypeptide having the sequence of amino acids 613-652 of SEQ ID NO:2, wherein the isolated polypeptide is a cysteine-rich domain.

20. An isolated polypeptide comprising a polypeptide having the sequence of amino acid 193-509 of SEQ ID NO:2, wherein the isolated polypeptide is a ZPH domain polypeptide.

21. The isolated polypeptide of any of claims 1, 2, 4, 19, or 20, wherein the isolated polypeptide consists of a functional fragment or variant.

22. An isolated polypeptide comprising a polypeptide selected from the group consisting of a polypeptide having the sequence of SEQ ID NO:7, a polypeptide having the sequence of SEQ ID NO:9, and a polypeptide having the sequence of SEQ ID NO:11.

23. The isolated polypeptide of claim 22, wherein the polypeptide consists essentially of a polypeptide selected from the group consisting of a polypeptide having the sequence of SEQ ID NO:7, a polypeptide having the sequence of SEQ ID NO:9, and a polypeptide having the sequence of SEQ ID NO:11.

24. An isolated polypeptide which binds selectively a polypeptide encoded by the isolated nucleic acid molecule of claim 1, 2, 4, or 14, provided that the isolated polypeptide is not PTPL1.

25. The isolated polypeptide of claim 24, wherein the isolated polypeptide binds to a polypeptide comprising the sequence of amino acids 658-898 of SEQ ID NO:2.

26. The isolated polypeptide of claim 24, wherein the isolated polypeptide binds to a polypeptide selected from the group consisting of a polypeptide comprising the sequence of SEQ ID NO:7, a polypeptide comprising the sequence of SEQ ID NO:9, and a polypeptide comprising the sequence of SEQ ID NO:11.

27. The isolated polypeptide of claim 24, wherein the isolated polypeptide binds to a polypeptide consisting essentially of the sequence of SEQ ID NO:2.

28. The isolated polypeptide of claim 24, wherein the isolated polypeptide is an antibody fragment selected from the group consisting of a Fab fragment, a $F(ab)_2$ fragment or a fragment including a CDR3 region selective for a PARG polypeptide.

29. An isolated complex of polypeptides comprising:

a polypeptide comprising the amino acid sequence of SEQ ID NO:12 bound to a polypeptide as claimed in claim 1, wherein the complex has phosphatase and GTPase activating activities.

30. The isolated complex of polypeptides of claim 29, wherein the polypeptides consist of the polypeptide of SEQ ID NO:12 and the polypeptide of SEQ ID NO:2.

31. A method for reducing Rho family signal transduction in a mammalian cell, comprising administering to the mammalian cell an amount of the isolated polypeptide of claim 17 effective to reduce Rho family signal transduction in the mammalian cell.

32. A method for reducing Rho family signal transduction in a mammalian cell, comprising administering to the mammalian cell an amount of the isolated polypeptide complex of claim 29 effective to reduce Rho family signal transduction in the mammalian cell.

33. A method for reducing the proliferation of a cancer cell, the proliferation of which is increased by Rho family protein signal transduction, comprising

administering to the cancer cell an amount of a polypeptide comprising a polypeptide encoded by the nucleic acid of SEQ ID NO:1 effective to reduce proliferation of the cancer cell.

34. A method of increasing Rho family protein signal transduction in a cell comprising

administering to the cell an amount of a dominant-negative variant of the polypeptide of SEQ ID NO:2 effective to increase the Rho family protein signal transduction in the cell.

35. The method of claim 34, wherein the dominant negative polypeptide has an inactivated GTPase-activating domain.

36. The method of claim 35, wherein the inactivated GTPase-activating domain is deleted.

37. The method of claim 35, wherein the inactivated GTPase-activating domain contains at least one inactivating point mutation.

38. A method of reducing binding of a protein which includes a PDZ4 domain to a protein which includes a PDZ4 domain binding site in a mammalian cell, comprising

contacting the mammalian cell with an agent which binds to the PDZ4 domain for a time effective to reduce binding of the protein which includes a PDZ4 domain to the protein which includes a PDZ4 domain binding site.

39. The method of claim 38, wherein the agent is a peptide comprising [at its carboxyl terminus] the amino acid sequence of SEQ ID NO:7, or conservative amino acid substitutions thereof excepting the terminal valine.

40. The method of claim 39, wherein the agent is a peptide selected from the group consisting of a peptide having the amino acid sequence of SEQ ID NO:7, a peptide having the amino acid sequence of SEQ ID NO:9 and a peptide having the amino acid sequence of SEQ ID NO:11.

41. The method of claim 38, wherein the agent is an antibody which binds to the PDZ4 domain.

42 The method of claim 41, wherein the antibody is a monoclonal antibody.

43. A method of modulating mast cell secretion in a subject, comprising:

administering to the subject in need of such treatment an amount of a modulator of PARG GTPase activating activity effective to modulate mast cell secretion.

44. A composition comprising:

the polypeptide of claim 17, and

a pharmaceutically acceptable carrier.

45. A composition comprising:

the polypeptide complex of claim 29, and

a pharmaceutically acceptable carrier.

46. A composition comprising:

the agent of claim 39, and

a pharmaceutically acceptable carrier.

47. A method for decreasing PARG GTPase activity in a subject comprising administering to a subject in need of such treatment an agent that selectively binds to an isolated nucleic acid molecule of claim 1 or an expression product thereof, in an amount effective to decrease PARG GTPase activity in the subject.

48. The method of claim 47, wherein the agent is an antisense nucleic acid.

49. The method of claim 47, wherein the agent is a polypeptide.

50. A method for identifying lead compounds for a pharmacological agent useful in the diagnosis or treatment of disease associated with PARG GTPase activating activity, comprising

forming a mixture comprising a PARG polypeptide or unique fragment thereof containing a PARG GTPase activating domain, a protein containing a GTPase specifically activatable by the PARG GTPase activating domain, and a candidate pharmacological agent,

incubating the mixture under conditions which, in the absence of the candidate pharmacological agent, permit a first amount of specific activation of the GTPase by the PARG GTPase activating domain, and

detecting a test amount of the specific activation of the GTPase by the PARG GTPase activating domain, wherein reduction of the test amount of specific activation relative to the first amount of specific activation indicates that the candidate pharmacological agent is a lead compound for a pharmacological agent which disrupts the GTPase activation activity of PARG, and wherein increase of the test amount of specific activation relative to the first amount of specific activation indicates that the candidate pharmacological agent is a lead compound for a pharmacological agent which enhances the GTPase activation activity of PARG.

51. The method of claim 50, wherein the GTPase specifically activatable by the PARG GTPase activating domain is Rho.

52. A method for identifying lead compounds for a pharmacological agent useful in the diagnosis or treatment of disease associated with PARG binding to a protein containing a PDZ4 domain, comprising

forming a mixture comprising a PARG polypeptide or unique fragment thereof containing a PDZ4 domain binding site, a protein containing a PDZ4 domain which selectively binds the PARG PDZ4 domain binding site, and a candidate pharmacological agent,

incubating the mixture under conditions which, in the absence of the candidate pharmacological agent, permit a first amount of selective binding of the protein containing a PDZ4 domain by the PDZ4 domain binding site, and

detecting a test amount of selective binding of the protein containing a PDZ4 domain by the PDZ4 domain binding site, wherein reduction of the test amount of selective binding relative to the first amount of selective binding indicates that the candidate pharmacological agent is a lead compound for a pharmacological agent which disrupts selective binding of a protein containing a PDZ4 domain by a PARG polypeptide containing a PDZ4 domain binding site and wherein increase of the test amount of selective binding relative to the first amount of selective binding indicates that the candidate pharmacological agent is a lead compound for a pharmacological agent which enhances selective binding of a protein containing a PDZ4 domain by a PARG polypeptide containing a PDZ4 domain binding site.

53. The method of claim 52 wherein the protein containing a PDZ4 domain is PTPL1.

Fig. 1

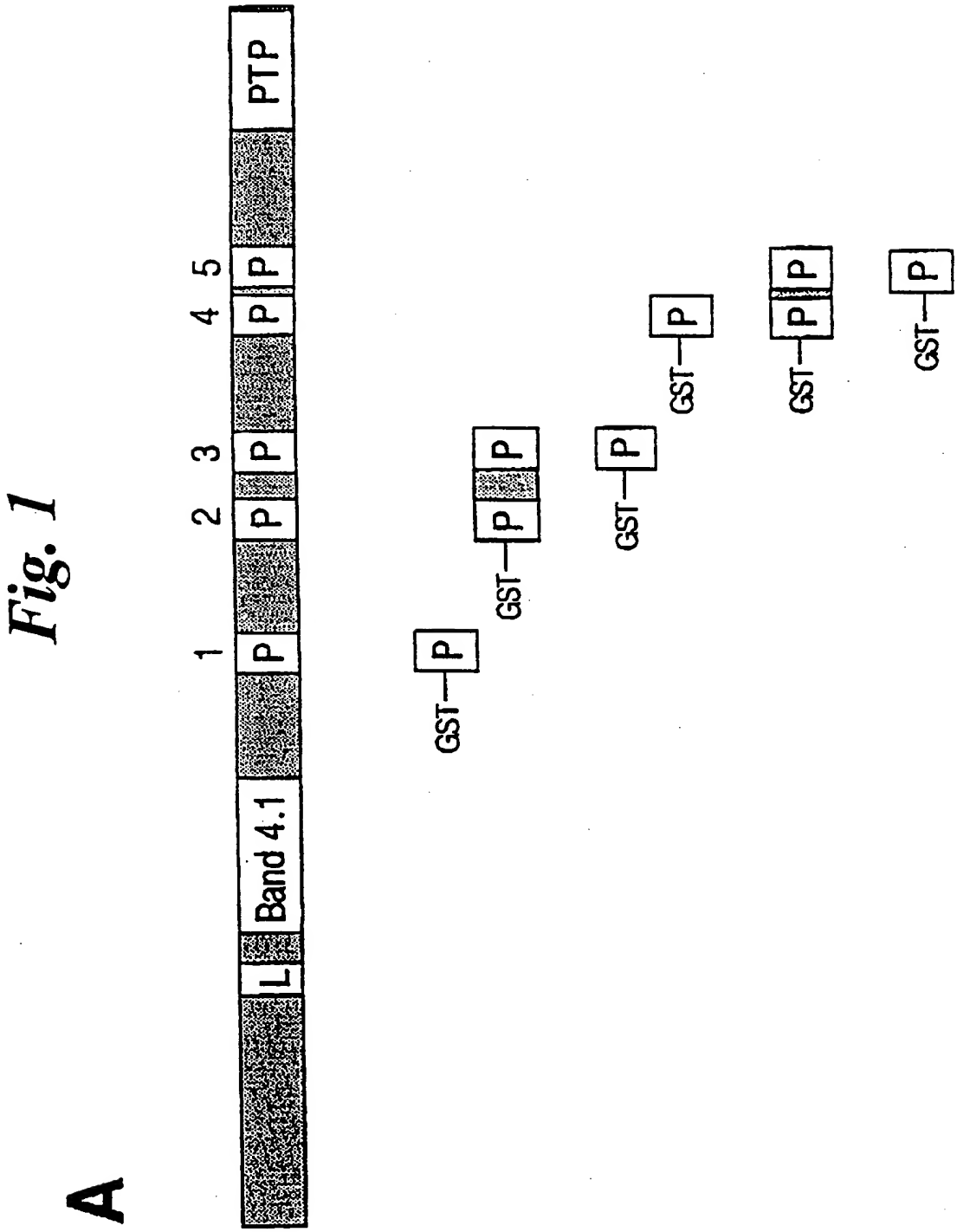


Fig. 1

Fig. 2

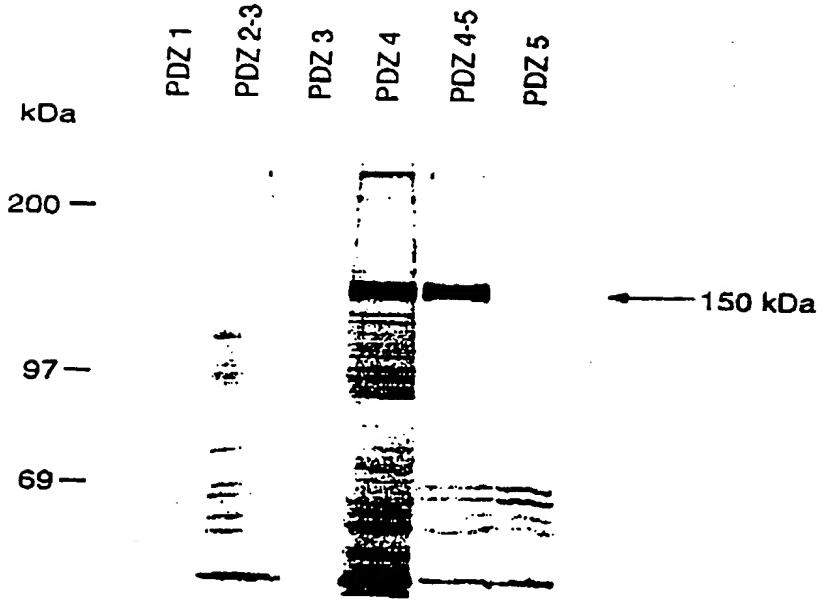


Fig. 3

A

```
MIAHKQKKTKKKRAWASGQLSTDITTSEMGLKSLSSNSIFDPDYIKELVNDIRKFSHILLYLKEAIFSDCFKEVIHIRLE      80
ELLRVLKSIMNKHQNLNSVDLQNAAEMLTAKVKAVNFTEVNEENKNDLFQEVFSSIETLAFTFGNILTNFLMGDVGNDSF     160
LRLPVSRETKSFENVSVESVDSSSEKGNFSPLELDNVLLKNTDSIELALSYAKTWSKYTKNIVSWVEKKLNLELESTRNM     240
VKLAEATRTNIGIQEFMPLQSLFTNALLNDIESSHLLQQTIAALQANKFVQPLLGRKNEMEKQRKEIKELWKQEQNKMLE     320
AENALKKAKLLCMQRQDEYEKAKSSMFRAEEEHLSSSGGLAKNLNKQLEKKRRLEEEALQKVEEADELYKVCVTNVEERR     400
NDVENTKREILAQLRTLVFQCDLTLKAVTVNLFHMQHLQAASLADRLQSLCGSAKLYQPGQEYSEFVKATNSTEEEKVDG     480
NVNKHLNSSQPSGFGPANSLEDVVRLPDSSNKIEEDRCSNSADITGPSFIRSWTFGMFSDSESTGGSSESRSLDSESISP     560
GQFHRKLPRTPSSGTMSSAQDLDEREPPSPSETGPNSLGTFKKTLMSKAALTHKFRKLRSPTKCRDCEGIVVFQGVECEE     640
CLLVCHRKCLENLVIICGHQKLPGKIHLFGAEFTLVAKKEPDGIPFILKICASEIENRALCLQGIYRVCGNKIKTEKLCL     720
ALENGMHLVDISEFSSHDICDVLKLYLRQLPEPFILFRLYKEFIDLAKEIQHVNEEQETKKNSLEDKKWPNMCIEINRIL     800
LKSKDLLRQLPASNFNSLHFLIVHLKRVVDHAEENKMNSKNLGVIFGPSLIRPRPQTAPITISSLAEYSNQARLVEFLIT     880
YSQKIFDGSLQPQDVMCSIGVVDQGCFPKPLLSPEERDIERSMKSLFFSSKEDIHTSESESKIFERATSFEESERKQNAL     960
GKCDACLSDKAQLLLDQEAESASQKIEDGKAPKPLSLKSDRSTNNVERHTPRTKIRPVSLPVDRLLLASPPNERNGRNMG    1040
NVNLDKFCKNPAFEGVNRKDAATTVCSKFNGFDQQTLQKIQDKQYEQNSLTAKTTMIMPSALQEKGVTTSLQISGDHSIN    1120
ATQPSKPYAEPVRSVREASERRSSDSYPLAPVRAPRTLQPQHWTTFYKPHAPIISIRGNEEKPASPSAACPPGTDHDPHG    1200
LVVKSMPDPDKASACPGQATGQPKEDSEELGLPDVNPMCQRPRLKRMQQFEDLEDEIPQFV*    1261
```

Fig. 3

B

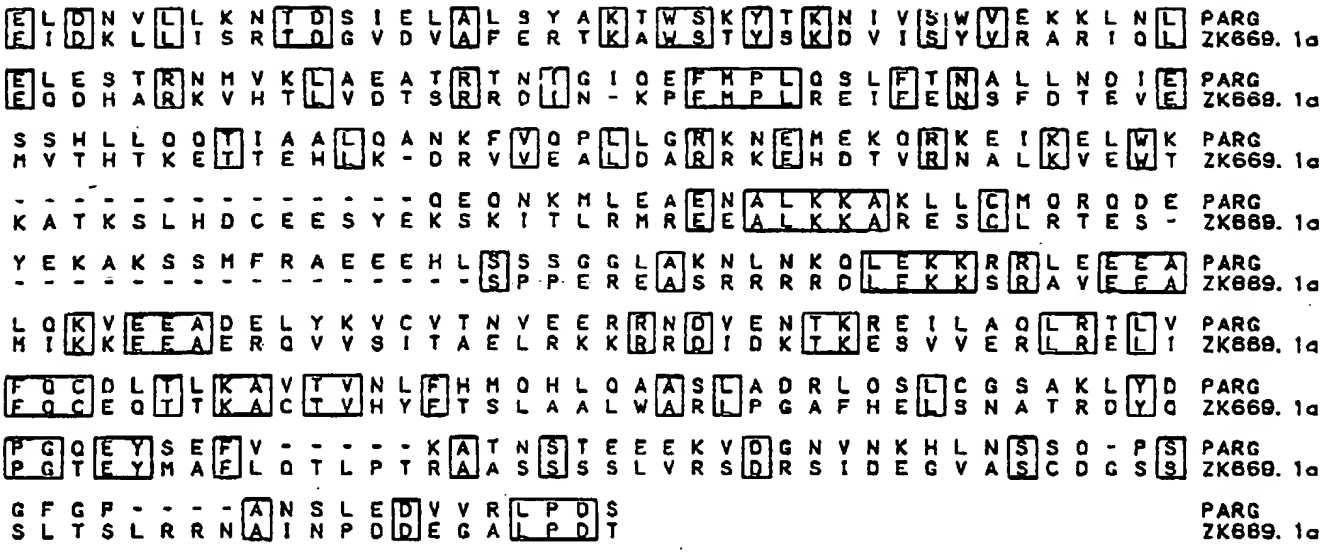


Fig. 3

C

PARG

ZK669.1a

Fig. 4

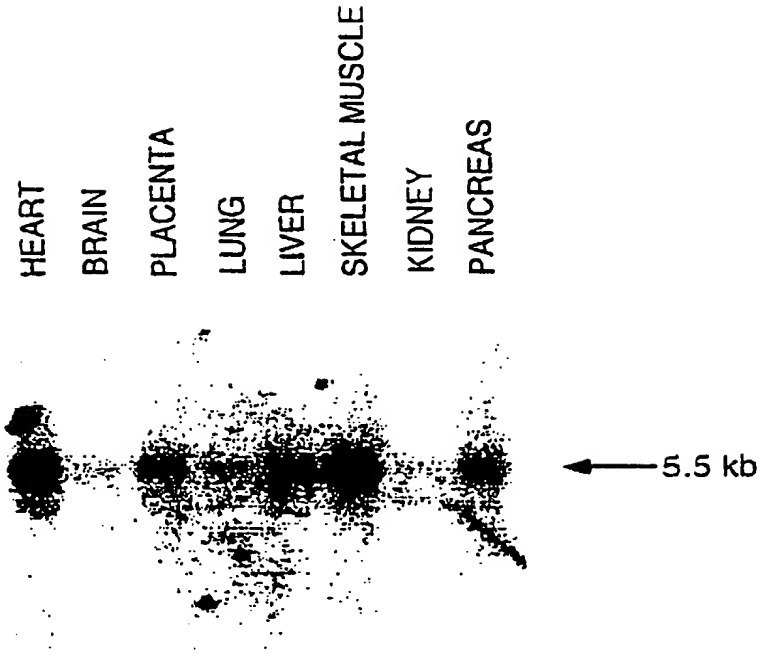


Fig. 5

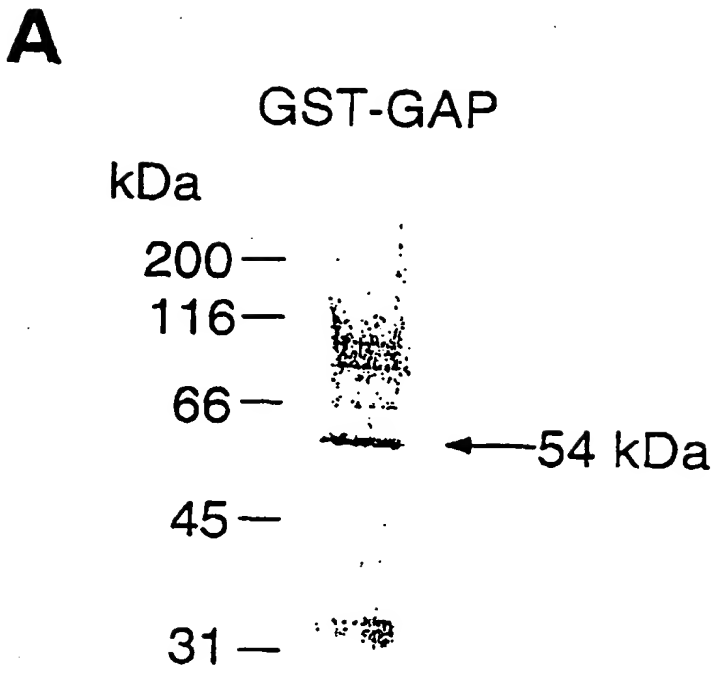


Fig. 5

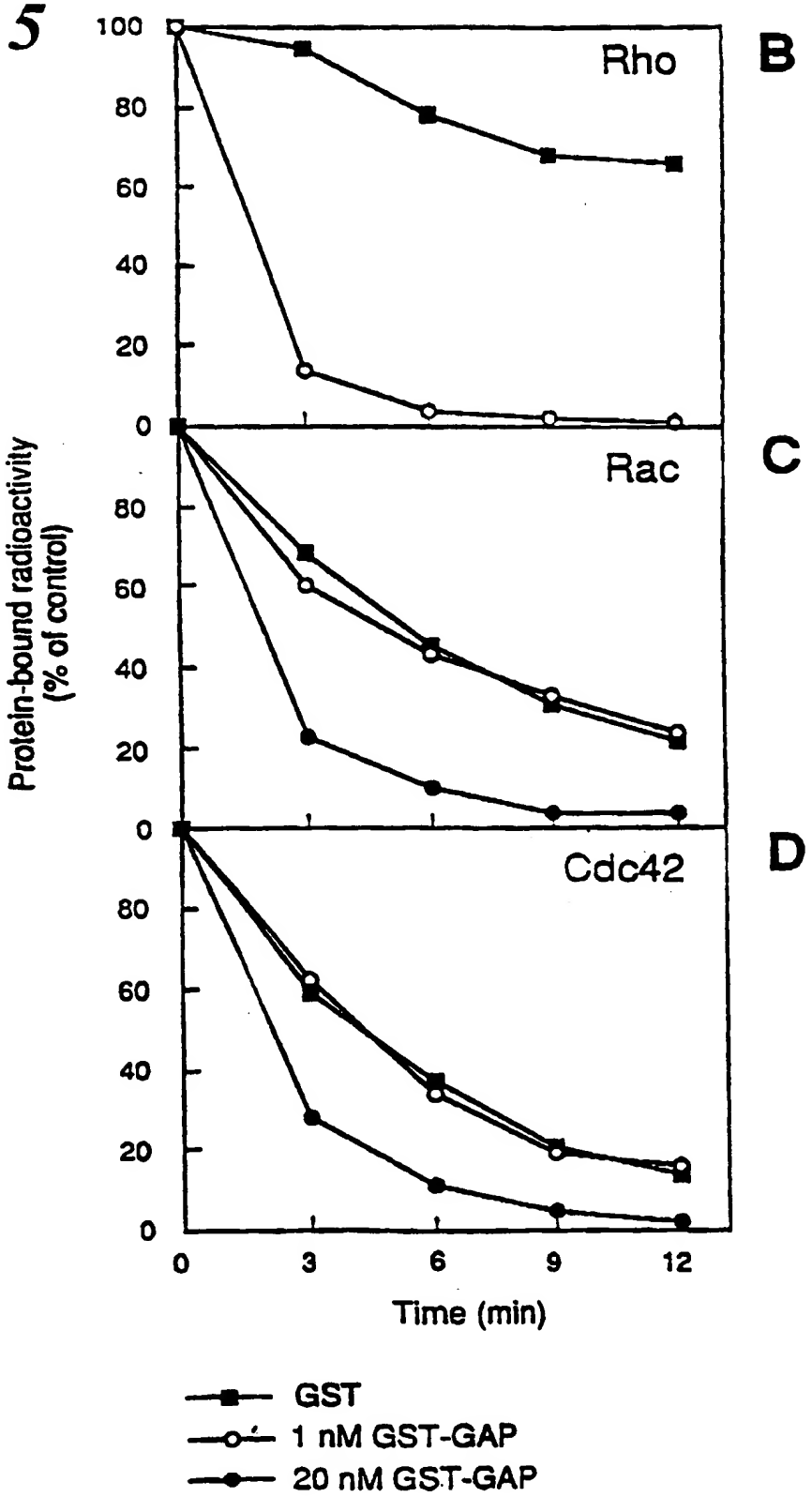
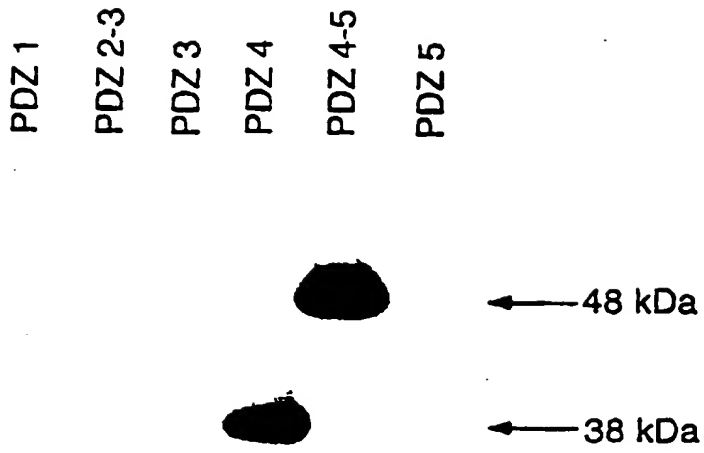


Fig. 6



INTERNATIONAL SEARCH REPORT

Internal Application No
PCT/US 98/03323

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/12 C07K14/47 C12N1/21 C12N5/10 C12N9/16
A61K38/17 A61K39/395 A61K38/46 A61K31/70 G01N33/68
C12Q1/44 C12Q1/42

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K A61K G01N C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	SARAS J. AND HELDIN C.-H.: "PDZ domains bind carboxy-terminal sequences of target proteins" TIBS TRENDS IN BIOCHEMICAL SCIENCES, vol. 21, no. 12, December 1996, pages 455-458, XP004063028	5-7, 13-30, 38-40, 44-46
Y	see page 455; table 1	31-33, 41-43, 50-53
A	see page 456; figure 1H see page 457, middle column - left-hand column	34-37, 47-49

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
13 July 1998	28.07.98

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Macchia, G

3

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
Y	LAMARCHE N. AND HALL A.: "GAPs for rho-related GTPases" TRENDS IN GENETICS, vol. 10, no. 12, December 1994, pages 436-440, XP002070373 cited in the application	31,32
A	see page 438, right-hand column, paragraph 5 - page 440, left-hand column, paragraph 1	34-37
Y	DIEKMANN D. AND HALL A.: "In vitro binding assay for interactions of Rho and Rac with GTPase-activating proteins and effectors" METHODS IN ENZIMOLOGY, vol. 256, no. Part B, 1995, pages 207-215, XP002070374 see page 207 see page 209, paragraph 2-3 see page 212, paragraph 3 - page 215	33,50,51
Y	PRICE L.S. ET AL.: "The small GTPases Rac and Rho as regulators of secretion in mast cells" CURRENT BIOLOGY, vol. 5, no. 1, 1 January 1995, pages 68-73, XP002070377 cited in the application see abstract see page 68, right-hand column, paragraph 2 - page 69, left-hand column, paragraph 1	33,43
Y	WO 95 06735 A (THE LUDWIG INSTITUTE FOR CANCER RESEARCH; GONEZ L.J.; SARAS J. (SE)) 9 March 1995 cited in the application	41,42, 52,53
A	see abstract see page 9, paragraph 3-4; figures 2,3 see page 23 see page 36, paragraph 2-3 see page 43 - page 44 see page 50; example 5 see page 84; claims 12,13	29,30
X	Database EMBL, entry EMEST7:HS763831, Accession number L49573 31 December 1995 98% identity with Seq.ID:1 nt.2722-3073 XP002070379 see the whole document	1,5-13, 21

-/--

3

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	Database EMBL, entry EMEST6:HS49515, Accession number T32495 13 January 1995 99% identity with Seq.ID:1 nt.3371-3671 XP002070381 see the whole document	1,5-13, 21
X	Database EMBL, entry EMEST10:HSC19G091, Accession number Z43348 6 November 1994 98% identity with Seq.ID:1 nt.3442-3749 XP002070382 see the whole document	1,5-13, 21
P,X	SARAS J. ET AL.: "A novel GTPase-activating protein for Rho interacts with a PDZ domain of the protein-tyrosine phosphatase PTPL1" THE JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 272, no. 39, 26 September 1997, pages 24333-24338, XP002070376 see the whole document	1-23, 29-32, 44,45
A	SYMONS M.: "The Rac and Rho pathways as a source of drug targets for Ras-mediated malignancies" CURRENT OPINION IN BIOTECHNOLOGY, vol. 6, no. 6, December 1995, pages 668-674, XP002071217	33
A	WHITE J.R. ET AL.: "Phosphorothioate-capped antisense oligonucleotides to Ras GAP inhibit cell proliferation and trigger apoptosis but fail to downregulate GAP gene expression" BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, vol. 227, 1996, pages 118-124, XP002070378 see abstract	47,48
A	WO 94 16069 A (SCHERING CORPORATION (US); NAKAFUKU MASATO (JP); KAZIRO YOSHITO (JP)) 21 July 1994 see abstract see page 3 - page 5 see page 18, line 28 - page 19, line 10 see page 79 - page 84; claims	31-33, 50,51
A	US RE35171 E (MCCORMICK FRANCIS P. ET AL.) 5 March 1996 see abstract see column 10, line 37 - column 11, line 25	50,51

-/--

3

INTERNATIONAL SEARCH REPORT

Inter. .onal Application No
PCT/US 98/03323

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	ASPENSTRÖM P. AND OLSON M.F.: "Yeast two-hybrid system to detect protein-protein interactions with Rho GTPases" METHODS IN ENZYMOLOGY, vol. 256, no. Part B, 1995, pages 228-241, XP002070375 see page 228 - page 229, line 4 see page 240, line 22 - page 241 ---	50,51
A	BOGUSKI M.S. AND MCCORNICK F.: "Proteins regulating Ras and its relatives" NATURE, vol. 366, 16 December 1993, pages 643-654, XP002057779 -----	

3

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 98/03323

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
Remark: although claims 31-42, as far as in vivo methods are concerned, and 43, 47-49 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

Inte... ...ional Application No
PCT/US 98/03323

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9506735 A	09-03-1995	AU 683299 B	06-11-1997
		AU 7644394 A	22-03-1995
		CA 2170515 A	09-03-1995
		EP 0789771 A	20-08-1997
		JP 9510861 T	04-11-1997
WO 9416069 A	21-07-1994	AU 6083894 A	15-08-1994
		CA 2153486 A	21-07-1994
		EP 0679185 A	02-11-1995
		JP 8507204 T	06-08-1996
US RE35171 E	05-03-1996	US 5104975 A	14-04-1992
		AT 156518 T	15-08-1997
		AU 627764 B	03-09-1992
		AU 4034989 A	05-02-1990
		DE 68928242 D	11-09-1997
		DE 68928242 T	18-12-1997
		EP 0466688 A	22-01-1992
		EP 0649908 A	26-04-1995
		WO 9000607 A	25-01-1990
		US 5234839 A	10-08-1993